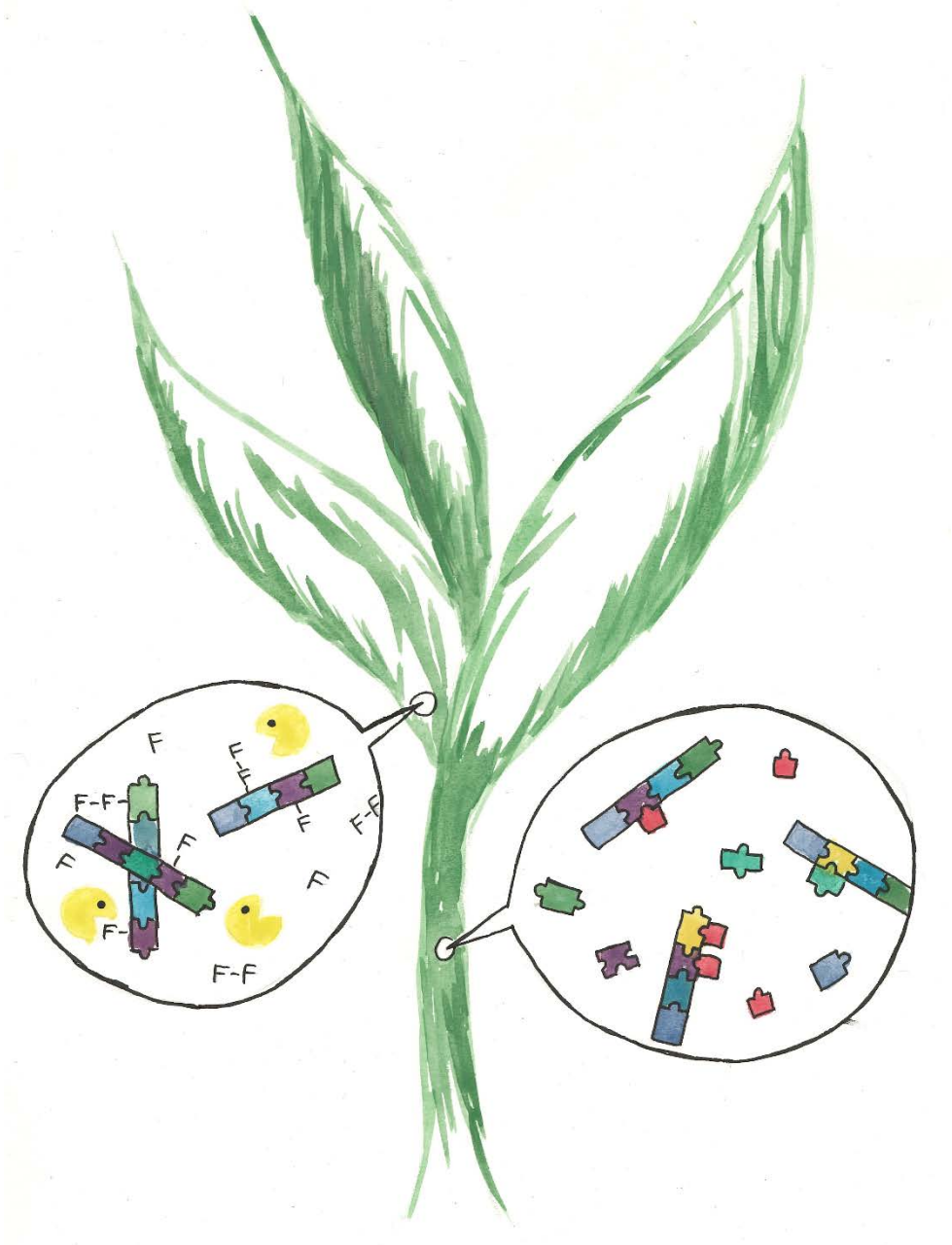
Synthesis of biomass-containing xylan fragments and evaluation of ferulic acid esterase activity

Underlin, Emilie Nørmølle

Publication date:
2018

Document Version
Publisher's PDF, also known as Version of record

Synthesis of biomass-containing xylan fragments and evaluation of ferulic acid esterase activity



Ph.D. thesis by Emilie Nørmølle Underlin, M.Sc.

November 2018

DTU Chemistry

Preface

The work presented in this thesis encompasses the results obtained during my Ph.D. studies from December 2015 to November 2018 at the Department of Chemistry, Technical University of Denmark as part of the Danish program to obtain a Ph.D. degree. The Ph.D. project was conducted under the supervision of Professor Robert Madsen as the main supervisor, and Professor Mads H. Clausen as co-supervisor. The project was financed by the Department of Chemistry, Technical University of Denmark.

A three month external stay was carried out at the Food Chemistry Department at Wageningen University and Research, Wageningen, The Netherlands, from April to early July 2018 under the supervision of Associate Professor Mirjam Kabel as part of the Ph.D. studies.

This dissertation covers four projects described in six chapters. Chapter one presents an introduction to xylan structures in biomass and the enzymatic breakdown of these as well as a presentation of previously published syntheses of xylan fragments. Chapter two first covers the development of a new glycosylation protocol as developed previously in the group for the synthesis of arabinoxylan fragments, which is followed by a description of the synthesis of three arabinoxylan fragments produced for this thesis. Chapter three first presents the previous efforts expended for the synthesis of glucuronoxylan fragments in the group, then the synthesis of two glucuronoxylans as produced for this thesis is presented. Chapter four covers the investigation into reactivity of ferulic acid esterases done during the external stay. Chapter five presents a short introduction to regioselective glycosylations followed by investigations of regioselective glycosylations catalyzed by molybdenum. Finally, chapter six provides an overall conclusion.

Emilie Nørmølle Underlin

November 2018, Kgs. Lyngby, Denmark

Acknowledgement

First and foremost, I would like to thank my supervisor Professor Robert Madsen for accepting me as a Ph.D. student and guidance during the last three years, which all made it possible for me to be a part of these projects.

I would also like to thank Professor Mads H. Clausen for being my co-supervisor and admitting me to the Bioorganic Groups Meetings, where I could discuss carbohydrate chemistry with other carbohydrate chemists.

Without the previous work of Ph.D. Maximilian Felix Böhm and Ph.D. Clotilde d'Errico, I would never have been able to get as far with these projects as I have or have been allowed to work with so many different and exciting topics during the last three years.

A great thank also goes to Associate Professor Mirjam Kabel for being my supervisor away from home and hosting me for my external stay at Wageningen University and Research. Furthermore I would have been lost without the kind and patient guidance of Ph.D. Matthias Frommhagen, who taught me the way around a biochemistry laboratory. Moreover, the project I worked on would not have been possible without the enzymes provided by Ph.D. Adiphol Dilokpimol and Professor Ronald P. de Vries at Utrecht University. Generally, I would also like to thank all the nice people at Food Chemistry for making my stay there both memorable and very enjoyable.

Laboratory work would not have been possible without the help of the technical staff of the chemistry department at DTU, hence thanks to sweet Anne Hector (who promised to stay until I left), Philip Charlie Johansen, Brian Brylle Dideriksen, and Lars Egede Bruhn. Many thanks also goes to the NMR team, Associate Professor Charlotte H. Gotfredsen, Professor Jens Ø. Duus, Kasper Enemark-Rasmussen, and again Anne Hector for maintaining the NMRs at all time. For help with MS measurements I have to thank Anders Holmgaard Hansen at DTU Biosustain.

The society does not function without money, so I thank DTU for financing my Ph.D., as well as the Oticon Fond, Tranes Fond, Jørgen Esmers Mindefond, and Knud Højgaards Fond for helping to finance my external stay.

The Madsen group, both former and present members, have made the last three years fun and enjoyable. Especially I need to thank Clotilde d'Errico, Carola Santilli, and Simone Samuelsen

for all the long and cosy talks about everything and sometimes absolutely nothing both in the laboratory and in the office.

Thanks to the fellow Ph.D. students and Postdocs at the organic department at DTU for providing a pleasant working environment. As special thanks goes to all the member of the "Pasta with Olives"-group, which originally was formed as a lunch group, but now has transformed into the most social group, which does not miss any chance for sharing food and good times.

Irene Boss – the first two years at DTU would not have been the same without you. Thanks for endless talks about chemistry and life both at home in Copenhagen and on vacations, and of cause for correcting this thesis. Without your comments and suggestions, this thesis would not have become what it is.

Thank you Fabrizio for always being ready with hugs and encouraging words if needed, for just being there and always believing in me.

Last but not least, a big thank you goes to my family for their support and love, for always picking up the phone for a talk even during dinnertime or welcoming me with open arms in Nibe or in Copenhagen.

Abstract

Lignocellulosic material from the plant cell wall would be an important and CO_2 neutral energy resource, if it could be properly utilised. However, intrinsic recalcitrance hampers the utilisation. Especially enzymatic degradation has received extensive attention as a way to solve this problem. An essential tool for development of new enzymes is well-defined model substrates, which can provide a deeper understanding of the substrate specificity and reactivity. Hence, chemically synthesised model substrates often outcompete isolated substrates from natural sources due to homogeneity, purity, and reproducibility.

The research presented in this thesis involved the synthesis of both arabinoxylan and glucuronoxylan fragments designed as model substrates for xylanases, α-arabinofuranosidases, and α-glucuronosidases. The syntheses involved the utilisation of thioxyloside building blocks linked in an iterative glycosylation strategy without pre-activation. The arabinose building block were synthesised through a small linear synthesis. On the other hand, the glucuronate building blocks were derived from the same thioglucopyranoside derivative through a divergent strategy (Figure 1).

One known class of the lignocellulosic enzymes are ferulic acid esterases. These enzymes are able to cleave the ester-linkages between monomeric or polymeric ferulic acids and a polysaccharide chain or lignin of the plant cell wall. The ferulic acid ester-linkages of the plant cell wall are known to hamper the degradation by other hemicellulosic enzymes. Hence, ferulic acid esterases are interesting tools for enzymatic degradation. The ferulic esterases can be classified into 13 subfamilies based on a phylogenetic three. In this thesis, 14 enzymes belonging to six different subfamilies were tested for specificity and reactivity towards both synthetic model substrates and natural substrates, thereby, being able to relate the reactivity and specificity to the classification.

Furthermore, the catalytic abilities of different molybdenum complexes for regioselective glycosylations have been tested.

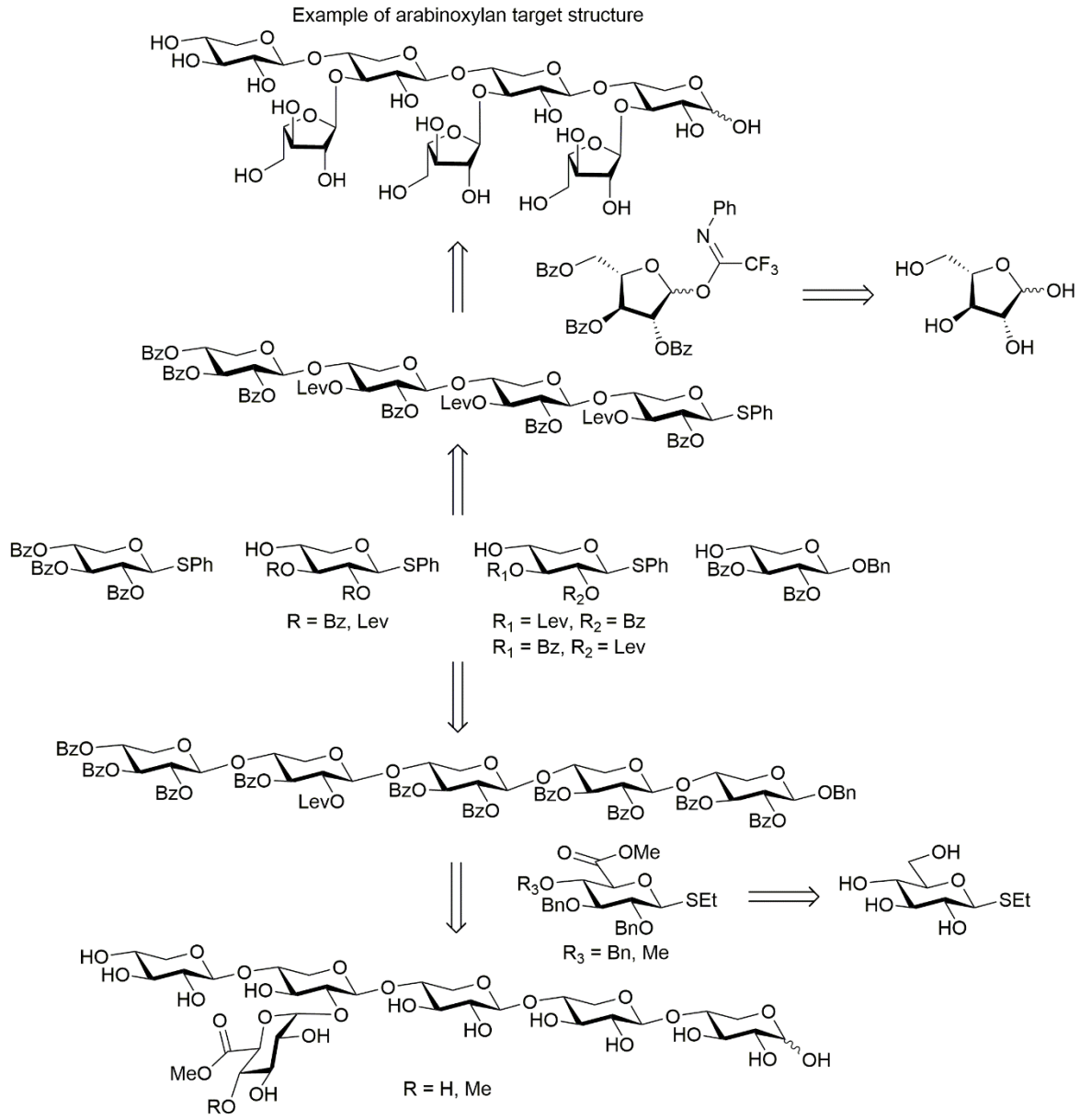


Figure 1: Retrosynthetic overview for the synthesis of arabino- and glucuronoxylan fragments.

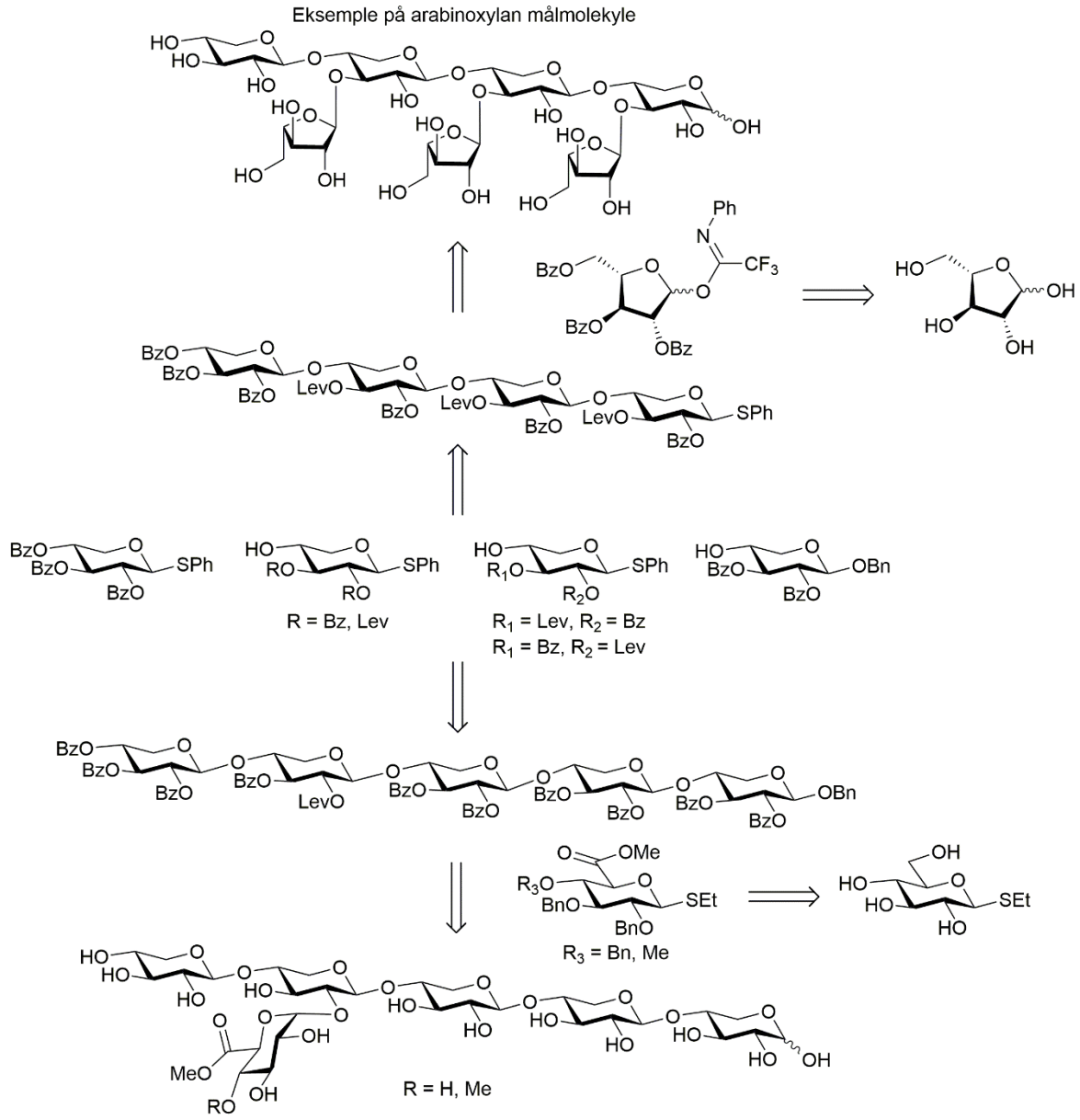
Resume

Lignocellulose materiale fra plante celle væggen ville være en vigtig og CO_2 neutral energiressource, hvis den kunne blive udnyttet ordentligt. Indbygget recalsistans hindrer dog udnyttelsen. Især enzymatisk nedbrydning har fået meget opmærksomhed som en mulig løsning på problemet. Et essentielt redskab for udviklingen af nye enzymer er veldefinerede model substrater, der kan give en dybere forståelse af substrat specificitet og reaktivitet. Derfor udkonkurrerer kemisk syntetiserede model substrater ofte isolerede substrater fra naturlige kilder i forhold til homogenitet, renhed og reproducerbarhed.

Forskningen præsenteret i denne afhandling omhandlede syntese af både arabinoxylan og glucuronanxylan fragmenter, der var designet som model substrater for xylanaser, α-arabinofuranosidaser og α-glucuronosidaser. Syntesen omfattede udnyttelsen af thioxylosid byggeblokke, der blev koblet via en iterativ glykosylerings strategi uden præaktivering. Arabinose byggeblokken blev syntiseret via en kortere lineær syntese. Modsat blev glucuronat byggeblokkene syntiseret via en divergent strategi fra det samme thioglucopyranosid (Figur 2).

Ferulasyre esteraser er en kendt klasse af lignocellulose enzymer. Disse enzymer kan kløve esterbindingen mellem monomeriske eller polymeriske ferulasyrer og en polysakkarid kæde eller lignin i plante celle væggen. Det er velkendt, at ferulasyre esterbindingerne i plante celle væggen hæmmer nedbrydningen via andre hemicellulose enzymer. Derfor er ferulasyre esteraser interessante redskaber for enzymatisk nedbrydning. Ferulasyre esteraser kan blive opdelt i 13 subfamilier baseret på et fylogenetisk træ. I denne afhandling blev 14 ferulasyre esteraser, der tilhører seks forskellige subfamilier, testet for deres specificitet og reaktivitet mod både syntetiske model substrater og naturlige substrater. Dermed kan specifiteten og reaktiviteten relateres til opdelingen.

Yderligere blev forskellige molybdenumkompleksers katalytiske evne for at dirigere regioselektive glykosyleringer testet.



Figur 2: Retrosyntetisk overblik for syntesen af arabino- og glucuronoxylan fragmenter.

Table of contents

Preface ... iii
Acknowledgement ... v
Abstract ... vii
Resume ... ix
Table of contents ... xi
List of abbreviations ... xv
1. Introduction ... 1
1.1 Lignocellulose structure ... 2
1.2 Xylans ... 3
1.2.1 Arabinoxylans ... 3
1.2.2 Glucuronoxylan ... 4
1.2.3 Glucuronoarabinoxylans ... 4
1.3 Ferulic acid and *p*-coumaric acid ... 5
1.4 Enzymatic degradation ... 8
1.4.1 Xylanases ... 10
1.4.2 α-arabinofuranosidases ... 11
1.4.3 α-glucuronidases ... 11
1.4.4 Ferulic acid esterases ... 11
1.5 Chemical synthesis of xylan oligosaccharides ... 15
1.5.1 Synthesis of homoxylans ... 18
1.5.2 Synthesis of heteroxylans ... 21
1.6 Glycosylations with glucuronic acid residues ... 26
2. Chemical synthesis of arabinoxylans fragments ... 31
2.1 Aim of the project ... 31
2.2 Optimisation of the glycosylation protocol ... 33
2.3 Synthesis of monomeric xylose building blocks ... 37
2.4 Synthesis of monomeric arabinose building block ... 40
2.5 Assembly of protected AX target structures ... 41
2.6 Deprotection of AX target structures ... 49
2.7 Future work ... 53

2.8 Concluding remarks 53
2.9 Experimental 55
3. Synthesis of glucuronoxylan fragments 87
3.1 Aim of the project 87
3.2 Previous results for the synthesis of GX fragments 88
3.2.1 Monomeric xylose building blocks 88
3.2.2 Monomeric glucuronic acid building blocks 91
3.2.3 Synthesis of GX target structures 93
3.3 Synthesis of a monomeric xylose building block 97
3.4 Synthesis of monomeric glucuronic acid building blocks 99
3.5 Assembly of protected glucuronoxylan target structures 101
3.6 Deprotection of GX targets 103
3.7 Future work 105
3.8 Concluding remarks 105
3.9 Experimental 107
4. Evaluation of ferulic acid esterase activity 127
4.1 Aim of the project 127
4.2 Ferulic acid esterases investigated in this project 128
4.3 Natural substrates tested 129
4.3.1 Sugar beet pectin 130
4.3.2 Corn fiber oligomers 130
4.3.3 Insoluble wheat arabinoxylans 130
4.3.4 Corn stover 131
4.3.5 Wheat straw 131
4.3.6 Corn stover lignin 131
4.4 Investigation of FAE specificity and reactivity 131
4.4.1 Determination of protein content in enzyme filtrates 131
4.4.2 FAE activity towards synthetic substrates 132
4.4.3 Type and quantification of hydroxycinnamic acids from the natural substrates 135
4.4.4 FAE activity towards natural substrates 139
4.5 Discussion of the observed FAE activities 143
4.6 Concluding remarks 147

4.7 Experimental 148
5. Regioselective glycosylations 153
5.1 Introduction to regioselective glycosylations 153
5.1.1 Tin-mediated glycosylations 154
5.1.2 Boron-mediated glycosylations 157
5.2 Possibility of molybdenum-mediated glycosylations 159
5.3 Studies of molybdenum-mediated glycosylations 160
5.3.1 Optimisation with glucose-based acceptor 160
5.3.2 Optimisation based on 1,2-*cis* diols 164
5.4 Concluding remarks 166
5.5 Experimental 167
6. Conclusion 177
7. Bibliography 179
Appendix 193

List of abbreviations

Ac	acetyl
acac	acetylacetonate
AGX	arabinoglucuronoxylan
APT	attached proton test (spectrum editing)
aq.	aqueous
Ar	aromatic
Ara*f*	arabinofuranosyl
Asp	aspartic acid
atm.	atmosphere
ax	axial
AX	arabinoxylans
Azmb	(azidomethyl)benzoyl
bs	broad singlet
BSA	bovine serum albumin
BSP	1-benzenesulfinyl piperidine
Bz	benzoyl
Calcd.	calculated
cat.	catalyst
Cbz	carboxybenzyl
CFoligo	corn fibre oligomers
ClAc	chloroacetyl
conc.	concentrated
C_q	quaternary carbon
CSA	camphorsulfonic acid
CS	corn stover
CSlignin	corn stover lignin isolate
d	doublet
Da	dalton
DBU	1,8-Diazabicyclo[5.4.0]undec-7-ene
DCC	*N,N'*-dicyclohexylcarbodiimide
DCE	1,2-dichloroethane
dd	doublets of doublets
ddd	doublets of doublets of doublets
DDQ	2,3-dichloro-5,6-dicyano-benzoquinone
diFA	di-ferulic acid
DMAP	4-(dimethylamino)-pyridine
DMAPA	3-(dimethylamino)-1-propylamine
DME	dimethoxyethane
DMF	dimethylformamide
DMSO	dimethyl sulfoxide

DMTST	dimethyl(methylthio)-sulfonium trifluoro-methanesulfonate
dq	doublets of quartets
DQF-COSY	double-quantum filtered correlation spectroscopy
dt	doublets of triplets
ELS	evaporative light scattering
eq	equatorial
equiv.	equivalent
ESI	electron spray ionization
Et	ethyl
FA	ferulic acid
FAE	ferulic acid esterase
GAX	glurunoarabinoxylan
gem	geminal
GH	glycosyl hydrolase
GlcA	glucuronic acid
Glu	glutamic acid
GX	glucuronoxylan
h	hour
H2BC	heteronuclear two-bond correlation
His	histidine
HMBC	heteronuclear multiple-bond correlation
HMPA	hexamethylphosphoroamide
HPLC	high performance liquid chromatography
HRMS	high resolution mass spectrometry
HSQC	heteronuclear single-quantum correlation
HSQC-TOCSY	heteronuclear single-quantum correlation- total correlation spectroscopy
m	multiplet
MBz	4-methyl-benzoyl
Me	methyl
MeGlcA	4-OMe glucuronic acid
MF	methyl ferulate
MilliQ	Millipore Corporation ultra pure water
min	minutes
MpC	methyl *p*-coumarate
MS	molecular sieves/mass spectrometry
MS^2	tandem mass spectrometry
Nap	2-naphthylmethyl

NBS	*N*-bromosuccinimide
NIS	*N*-iodosuccinimide
NMR	nuclear magnetic resonance
p	*para*
pCA	*p*-cumaric acid
Ph	phenyl
PMB	*para*-methoxyphenyl
ppm	parts per million
*p*TSA	*para*-toluenesulfonic acid
q	quartet
R_f	retention factor
RP	reverse phase
R_t	retention time
r.t.	room temperature
s	singlet
sat.	saturated
SBP	sugar beet pulp
Ser	serine
SF	subfamily
t	triplet
TBAB	tetrabutylammonium bromide
TBAI	tetrabutylammonium iodide
TBS	*tert*-butyldimethylsilyl
TBSOTf	*tert*-butyldimethylsilyl trifluoromethanesulfonate
td	triplets of doublets
TEMPO	2,2,6,6-tetramethyl-1-piperidinyloxy
TESOTf	triethylsilyl trifluoromethanesulfonate
Tf_2O	trifluoromethanesulfuric anhydride
TFA	trifluoroacetic acid
TFAA	trifluoroacetic anhydride
TfOH	trifluoromethanesulfonic acid
THF	tetrahydrofuran
TLC	thin layer chromatography
TMSOTf	Trimethylsilyl trifluoromethanesulfonate
TOCSY	total correlation spectroscopy
triFA	tri-ferulic acid
Trt	trityl
UHPLC	ultra-high performance liquid chromatography
UV	ultraviolet light
vol.%	volume percentages
WAX-i	insoluble wheat arabinoxylan
WS	wheat straw

w/w	weight per weight

1. Introduction

In the last decades the goal of becoming independent of fossil fuels and petro chemistry has become more prominent, and consequently an exponential increase in the global interest for renewable resources has been observed. Biomass and biomass-derivatives have been regarded as promising alternatives as these offer a steady supply of suitable organic carbon for production of fuels and fine chemicals while at the same time having the benefit of net zero carbon emmision.[1,2]

The most abundant and renewable biomass resource is lignocellulose, which is estimated to have an annual production of 10-50 billion ton.[3] Lignocellulose is defined as the non-eatable part of plants, and is also called dietary fibres. The accumulation of lignocellulosic material as forestry, agricultural, and agro-industrial waste is a further environmental problem. Hence, lignocellulose has tremendous potential as starting material for bioethanol, paper, and fine chemicals.[1,2]

However, the intrinsic recalcitrance to chemical, physical, and enzymatic degradation is hampering the utilisation of lignocellulose. Consequently, this has motivated extensive research to address this problem, especially with focus on enzymatic degradation, as this constitute a mild treatment with comparative low resource demand.[4,5] The enzymatic

hydrolysis is both affected by enzyme- and substrate-related factors, which reduces the accessibility to the targets. The enzyme-related factors are enzyme activity, product inhibition, stability, and synergism while the substrate-related factors focus on improving the accessibility by different pre-treatment methods as much as possible whilst still being economically viable. Due to the high cost and resource consumption of pre-treatment, new innovative and sustainable methods are needed, which especially include the development and discovery of novel enzymes needed for complete degradation of the lignocellulose components.[6–8]

Another aspect of lignocellulose is the functions as dietary fibre. Dietary fibres resist host digestion, absorption, and adsorption processes. They are fermented by intestinal microbes, and stimulate growth and/or activity of intestinal bacteria which are associated with health and wellbeing.[9] For e.g. swine and poultry there is a fine line between too high and too low levels of dietary fibres in the feed, as high dietary fibre contents are associated with decreased nutrient value and lower net energy utilisation of the feed. However, dietary fibres are also being extensively investigated as probiotics for growth. This is especially important since antibiotics have for a long time been utilised as growth promotors, however the regular use of antibiotics can lead to bacterial resistance both for human and animal diseases.[9–11]

1.1 Lignocellulose structure

The three major components of lignocellulose are cellulose, hemicellulose, and lignin.[2] Cellulose is the main component of the plant cell wall and it represents the most abundant organic polymer on Earth. It is an unbranched homo-polymer consisting of β-(1→4)-linked D-glucose units which can pack tightly in fibrils.[4] In contrast, hemicellulose has classically been described as a group of heterogeneous polymers including xyloglucans, xylans, mannans, glucomannans, and β-(1→3,1→4)-glucans, which can be extracted by an alkaline solution. However this definition is not useful, as some pectins (*vide infra*) also can be extracted by alkaline treatment, as well as some mixed-linked glucans in the grass cell wall can be readily extracted without alkaline treatment.[12] In 2010 H. V. Scheller and P. Ulvskov defined hemicellulose as "a group of wall polysaccharides that are neither cellulose nor pectin and by having β-(1→4)-linked backbones of glucose, mannose or xylose."[12] Moreover, hemicelluloses are generally characterised by being rich in substitutions. The main function of hemicelluloses are strengthening of the plant cell wall through cross-interaction with both cellulose and lignin.[12] Xylans are generally acknowledged as the largest group of the hemicelluloses and can account for up to 50% (w/w).[13] Xylans consists of β-(1→4)-D-xylose backbones, which are decorated

with arabinoses, glucuronic acids, xyloses, and acetates.[12] The main focus of the introduction in this thesis will be the structure, degradation, and chemical synthesis of xylans. The last lignocellulose component is lignin. This is an aromatic heteropolymer consisting of three phenylpropane monomers; coniferyl, sinapyl and *p*-coumaryl alcohol, which form a complex 3D-network. The main function of lignin is water transportation, pathogen defence and structural support.[14,15] The structural support highly relies on cross-linking to hemicellulose through ether- and ester-linkages comprising glucuronic acid and arabinosyl-ferulic acid substituents.[16,17]

The last polysaccharide polymer in the plant cell wall is pectin. This polymer is not characterised as being part of lignocellulose, however the structure should still be mentioned. Pectin provides cell adhesion, structural support, and defence responses against e.g. pathogens, and the most common commercial use is in the food industry as a gelling agent. It is a highly heterogeneous structure consisting of many different carbohydrates, which are arranged in covalently connected domains. The main domains are homogalacturonan, rhamnogalacturonan I and II, and/or xylogalacturonan, where also rhamnogalacturonan I can have ester-linkages to ferulic acid providing structural support.[18,19]

1.2 Xylans

Xylans are highly complex and heterogeneous structures differentiating in quantity, distribution, and branching units, which characterise the function and structure of the different plant cell walls in nature. Despite the high degree of variation it is generally accepted to divide the xylans into homoxylans and heteroxylans, where the heteroxylans include arabinoxylans (AX), glucuronoxylans (GX), and glucuronoarabinoxylans (GAX).[13] Apart from the saccharide substituents xylans are also acetylated in various degrees, mainly in the *O*-3 position and less often in the *O*-2 position.[12,13]

1.2.1 Arabinoxylans

Arabinoxylans are highly represented in the cell wall of cereal grains (e.g. wheat, rye, barley, oat, rice, and corn), and it is even the main hemicellulose component in wheat bran. Over the last decades AX have been found to be important dietary fibres with a positive impact on gut associated microbial growth and the immune system. AXs have a backbone of β-(1→4)-linked D-xylose with L-arabinofuranosyl (Ara*f*) residues attached. The xylan unit can be either single substituted with Ara*f* linked as α-(1→3) or disubstituted by α-(1→3) and α-(1→2)-linked Ara*f* units (Figure 1-1). The ratio of arabinose to xylose can range from 0.2 to 1.2. Additionally, the

O-5 position of Ara*f* residues can be bonded to ferulic acid (FA) and *p*-coumaric acid (pCA), which can give rise to cross-linking (this will be further discussed in paragraph 1.3).[12,13,20]

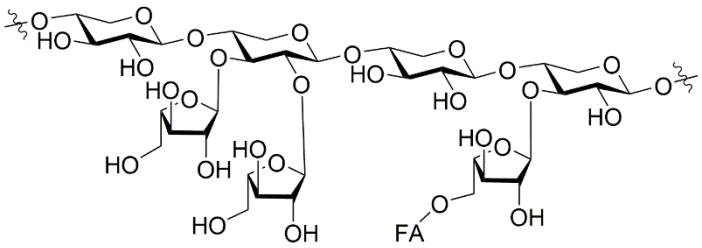


Figure 1-1: Model structure of arabinoxylan.

1.2.2 Glucuronoxylan

Glucuronoxylan is the main hemicellulose component of hardwoods. For these structures, the D-xylose backbone is substituted with α-D-glucuronic acids (GlcA) or its 4-methyl-form (MeGlcA), where the average substitution ratio is 1:10 of glucuronic acid to xylose units. Furthermore, GX are highly acetylated with up to seven acetyls per 10 xylose units (Figure 1-2).[13,19,21]

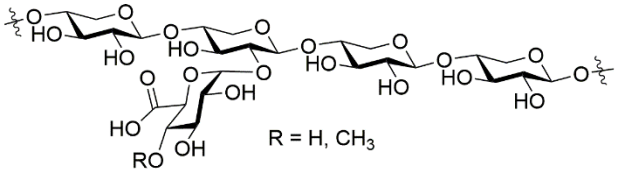


Figure 1-2: Model representation of glucuronoxylan.

In Kraft pulping, which is a preliminary step in paper production, utilisation of alkaline conditions convert wood into pulp, which mainly consists of cellulose. However, during this process the glucuronic acid can be partially converted into hexenuronic acid through β-elimination and 4-*O*-methyl-L-iduronic acid via *C*-5 epimerisation, which can interfere with the paper production process.[22,23]

1.2.3 Glucuronoarabinoxylans

Depending on the ratio of glucuronic acid and arabinose to xylose, a distinction in the name can be found, as either GAX or arabinoglucuronoxylan (AGX). Common for these structures are the D-xylose backbone, which can have single MeGlcA and L-arabinofuranosyl residues attached in the *O*-2 and *O*-3 position, respectively, as well as being acetylated. AGX is a major component in softwood, where the ratio of xylose to MeGlcA is 5-6:1, and only a small proportion of the backbone is substituted with Ara*f* (approximately 1.3 per 10 units of xylose). In comparison,

GAX has ten times fewer MeGlcA to Ara*f*, and some of the xylose units are even doubly substituted with Ara*f*. Furthermore, the Ara*f* units can for both structures be substituted with D-xylose as β-(1→2)- or β-(1→3)-linkages, and FA at the *O*-5 position (Figure 1-3).[13,24]

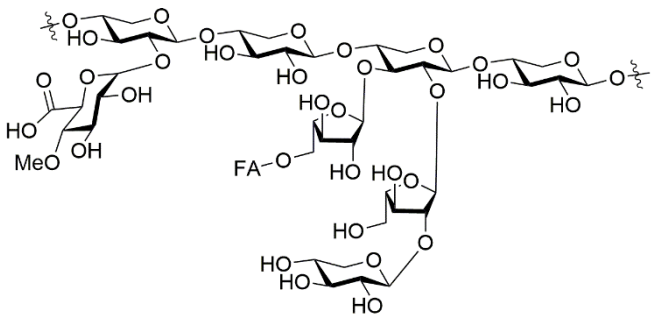


Figure 1-3: Model of a GAX molecule.

1.3 Ferulic acid and *p*-coumaric acid

As written previously the hydroxycinnamic acids ferulic acids and to a lesser extend *p*-coumaric acids (Figure 1-4) are molecules that are cross-linked to the different polymers of the lignocellulosic material especially in grasses.[25] They are found covalently linked to the saccharide sidechains in AX, GAX and the neutral sidechain in rhamnogalacturonan I in pectin. Furthermore, FA can form diferulic acids that can cross-link polymers. These esterifications and cross-links increase the intrinsic recalcitrance to saccharification and biodegradation of the plant material.[26,27] Ferulic acid in itself is furthermore an interesting molecule as it also have pharmaceutical functions, as well as applications in foods and cosmetics.[28]

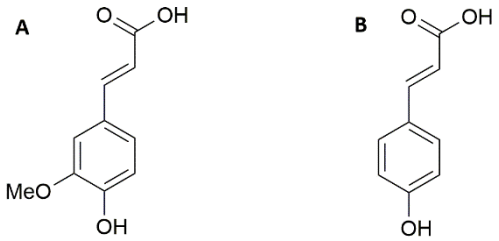


Figure 1-4: Chemical structure of A: FA and B: pCA.

In AX and GAX, ferulic acid is mainly linked through an ester-bond at the carboxylic acid of FA and the *O*-5 of Ara*f*. Additionally, the FA can be further linked to another FA molecule or lignin (Figure 1-5).[26,27]

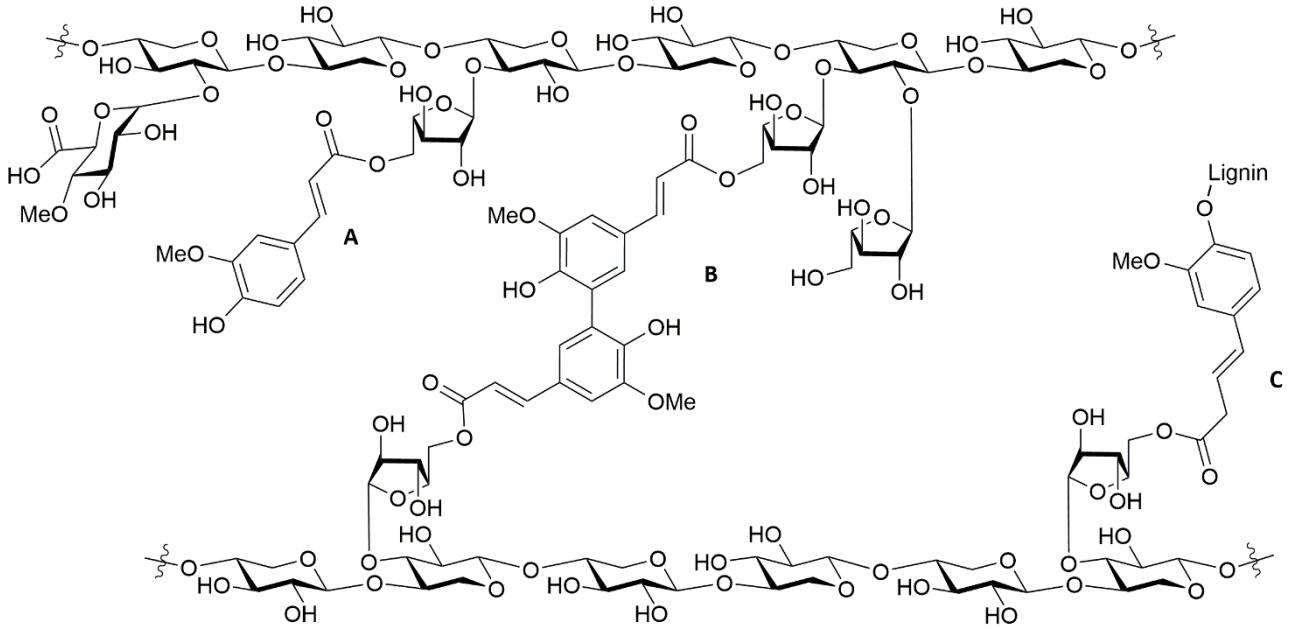


Figure 1-5: Schematic representation of the feruloylation of GAX with the backbone of β-(1→4)-D-xylopyranose. A) A feruloyl group attached to O-5 of an arabinofuranosyl residue. B) A diferuloyl cross-link at O-5 of the arabinofuranosyl residues, the dimer being a 5-5'. C) A feruloyl group attached to lignin and the O-5 of an arabinofuranosyl residue.[26,27]

Rhamnogalactoronan I in pectin consists of a backbone of α-L-rhamnose-(1→4)-α-D-galactoronan-(1→2). Additionally, the backbone is substituted with neutral sidechains based on arabinosyl and galactosyl moieties. The arabinosyl sidechain consists of α-(1→5)-linked Ara*f* units, and the galactosyl component of β-(1→4)-linked D-galactose chains. FA can be linked to these neutral sidechains, where the ester-link is between the carboxylic acid and the *O*-6 hydroxyl group on β-D-galactopyranose and the *O*-2 or *O*-5 hydroxyl of the α-L-arabinofuranosyl unit (Figure 1-6).[26,27]

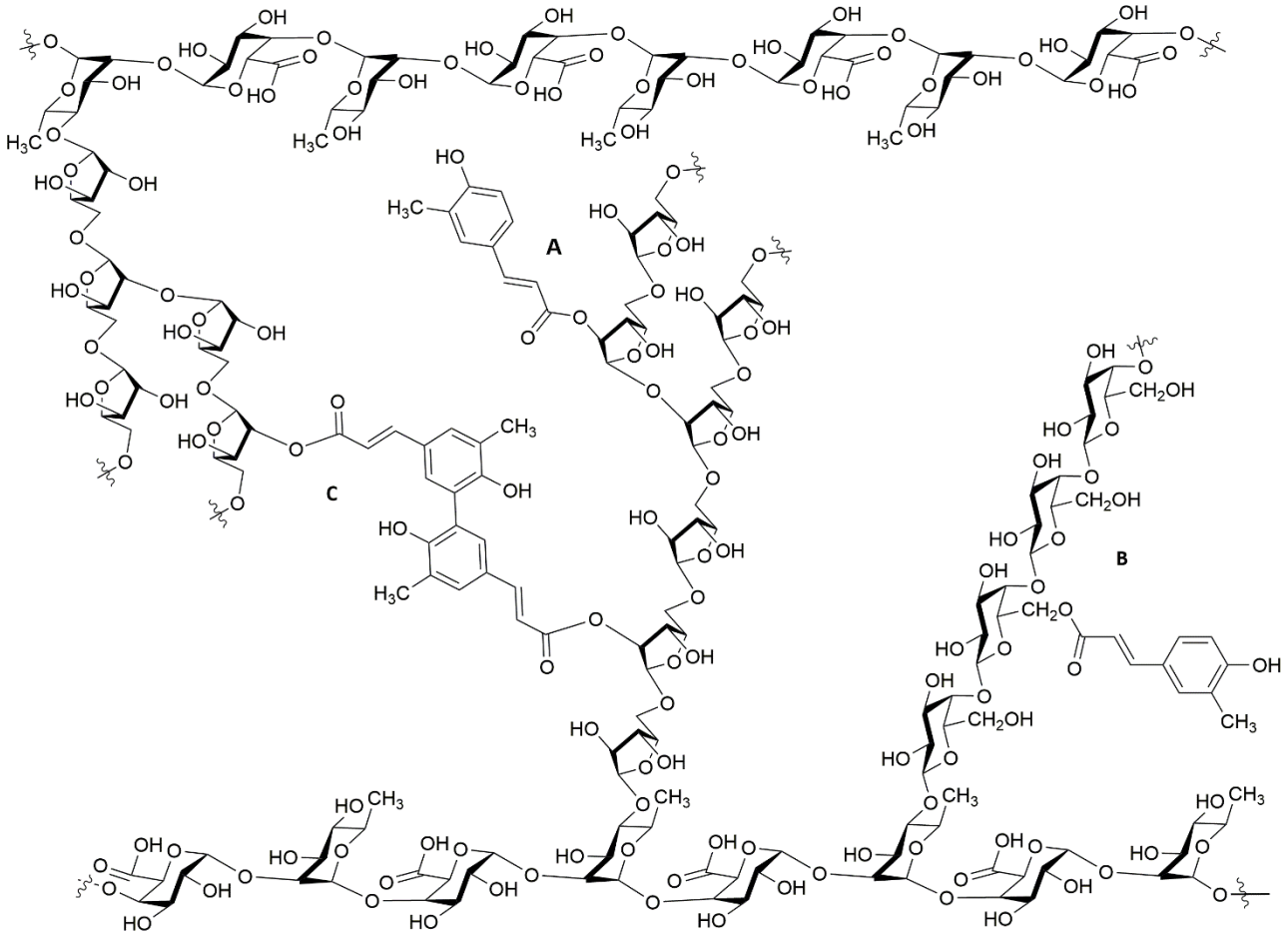


Figure 1-6: Schematic representation of feruloylation in rhamnogalactoronan I. A) FA attached to O-2 of an Ara*f* residue. B) A feruloyl group attached to the O-6 of a galactopyranosyl residue. C) A diferuloyl cross-link at O-2 of the Ara*f* residues, the dimer being a 5-5'.[26,27]

The diferulic acids (diFA) seen in Figure 1-5 and Figure 1-6 are examples of oxidative cross-links thereby connecting two polysaccharide backbones. Several different diFA linkages have been observed of which the main linkages can been seen in Figure 1-7.[26,27,29–31] Furthermore FA can mediate a covalent linkage from the polysaccharide to lignin as seen in Figure 1-5.[25,30,31] Different types of linkages of FA to lignin have been observed, where the β-*O*-4 as in Figure 1-5 is one of the most common.[32] Dimerization of FA requires close proximity of the FAs, hence diFA linkages have mainly been found in the high-arabinose substitution region of AX and GAX. Generally, a dramatic increase in mechanical strength is provided by the cross-linking of the cell wall polysaccharides and lignin as well as acting as a barrier towards invading microbes and hydrolytic enzymes.[25,33] Furthermore, several different triferulic acids (triFA) have been found in maize bran,[34–36] which could indicate the linkage of three polysaccharide chains. This expands the possibility for cell wall cross-linking and implies that the cell wall structure is possibly even more complex than previously assumed. pCA on the other hand is mainly γ-

esterified to terminal lignin residues and to a less extend to the polysaccharides.[32] However, overall it can be said that, saccharification and biodegradation of the plant cell wall is being further hampered by the presence of FA, pCA, and diFA cross-links.

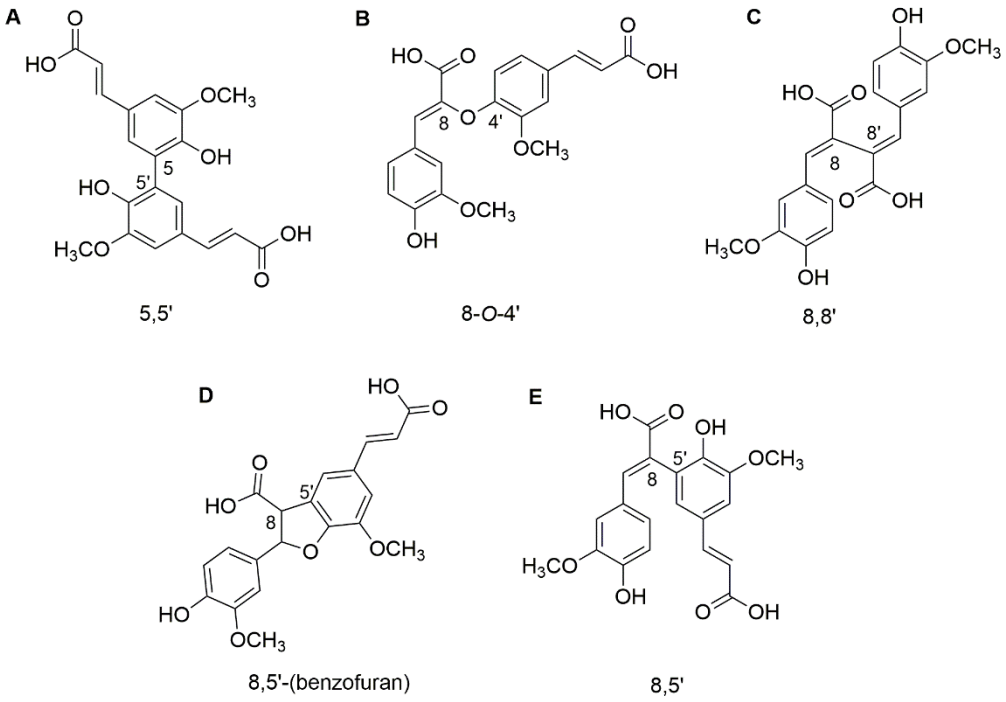


Figure 1-7: Chemical structure of different diFAs found in nature. A) 5,5'-diferulic acid. B) 8-O-4'-diferulic acid. C) 8,8'-diferulic acid. D) 8,5'-(benzofuran)-diferulic acid. E) 8,5'-diferulic acid.[27]

1.4 Enzymatic degradation

For lignocellulose to be a resourceful raw material for e.g. bioenergy production, the degradation has to be effective. Enzymatic degradation is especially interesting as it is environmentally friendly, specific, mild, and the production of by-products is low.[7] An envisioned solution is suitable enzyme cocktails with several different types of enzymes, all defined as hemicellulases, working in synergy to disrupt the complex lignocellulose structures (Figure 1-8). The combination being more effective than the sum of the individual enzymes, as they are working in a synergistic fashion, being defined as heterosynergy.[37,38] Additionally, ferulic acid esterases (FAE) can be added to further enhance the degradation.[39,40]

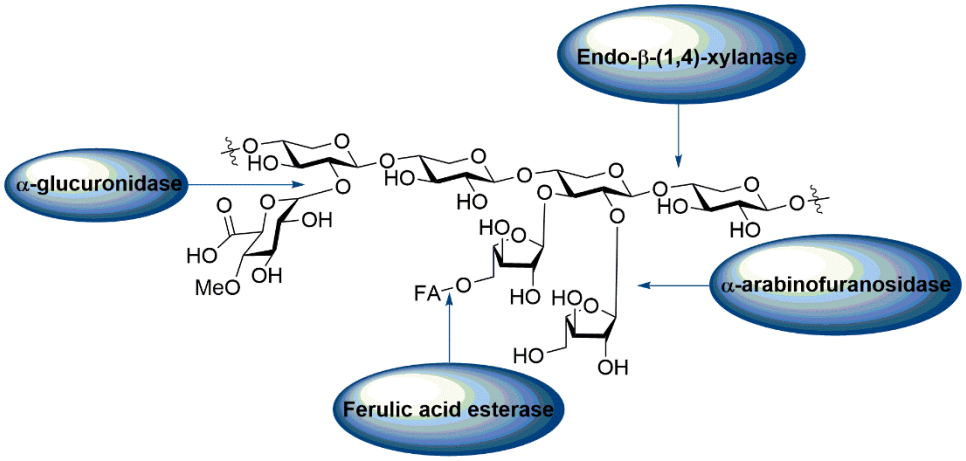


Figure 1-8: Schematic overview of the enzymes involved in the degradation of a substituted xylan chain with the active sites indicated.

Optimisation of the enzyme cocktails require a deep knowledge of the individual enzyme's specificity and synergy with others. For this purpose, the design of well-defined target structures of lignocellulose would be an indispensable tool to acquire an extended understanding of the enzymes and their possible applications. A general overview of the different enzyme classes involved in xylan degradation will be presented below with special focus on ferulic acid esterases.

Glycosidases can hydrolyse the glycosidic linkage either with inversion or retention of the stereochemistry at the anomeric carbon. The mechanism of retaining glycosidases involves two consecutive S_N2 reactions, where first a substitution is performed by a nucleophilic carboxylic acid in the active site aided by another carboxylic acid protonating the aglycone, the result being a covalent enzyme-substrate intermediate. In the second displacement, a water molecule acts as the nucleophile resulting in the release of the substrate with retention of the stereochemistry (Figure 1-9 A). For inverting glycosidases only a single S_N2 reaction takes place, where a carboxylic acid in the active site acts as a base activating a water molecule for attack at the anomeric carbon. The aglycone is being protonated by another carboxylic acid, leading to the release of the substrate with inversion of the stereochemistry (Figure 1-9 B).[37,41]

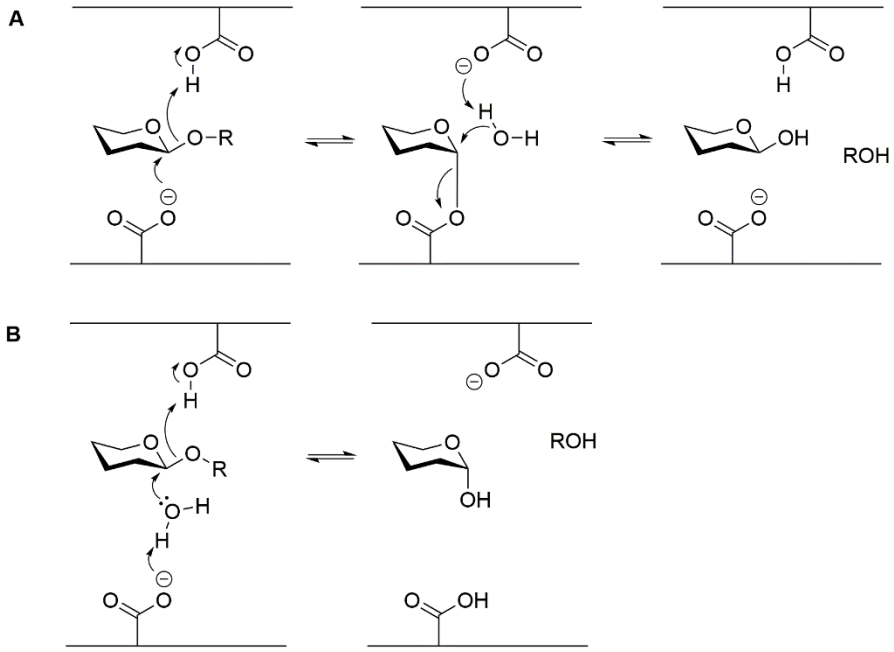


Figure 1-9: General mechanisms of glycosyl hydrolases as retaining glycosidases (A) or inverting glycosidases (B).

1.4.1 Xylanases

Xylanases are a diverse group of *O*-glycoside hydrolases which are able to degrade the β-(1→4)-linkages of xylans. Generally, xylanases are divided into endo-β-(1→4)-xylanases and β-xylosidases. The endo-xylanases can cleave internally on the xylan backbone chain, while β-xylosidases liberates a xylose monosaccharide by cleavage from the non-reducing end of a short oligosaccharide. The binding sites of the xylanases are termed subsites where the cleavage takes place between the -1 subsite (non-reducing end) and +1 (reducing end) (Figure 1-10).[42,43]

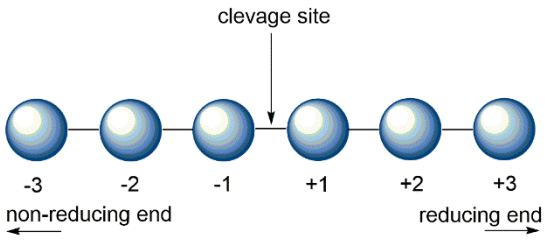


Figure 1-10: Schematic representation of the sugar binding site of hydrolases.

Endo-xylanases, which are retaining enzymes, have been found to belong to glycosyl hydrolase (GH) families 5, 7, 8, 10, 11, and 43. Generally, the activity of xylanases are hampered by substitution close to the binding side, although in recent years several GH families have been found to accept some substitutions close to the cleavage site. The most studied families of endo-xylanases are GH10 and 11 which have been found to accept either MeGlcA or Ara*f* in the +1

and +2 subsite, respectively.[44,45] Furthermore, major interest have been given to two other families of xylanases; GH5 and GH30. GH5 accommodates a *O*-3 linked Ara*f* in the -1 subsite[46,47], where GH30 can accommodate either a *O*-2 linked MeGlcA or GlcA at the terminal non-reducing end.[48,49] β-Xylosidases have been found to belong to GH3, 30, 39, 43, 52, 54, and 62. These perform the hydrolysis of the glycosidic bond with retention. The exceptions are GH43, which are inverting enzymes, and GH62 for which the mechanism has not yet been elucidated. As mentioned above xylosidases cleave the β-(1→4)-linkage next to the non-reducing end of short oligosaccharides.[37] Even though xylanases can depolymerise the xylan backbone, several different types of enzymes are needed to achieve debranching and hence full degradation.

1.4.2 α-arabinofuranosidases

Arabinofuranosidases are often termed accessory enzymes to xylanases, as these are needed for optimised function of xylanases. α-Arabinofuranosidases have been found in GH3, 43, 51, 54, and 62.[43] However, the major part of arabinofuranosidases belong to GH43 and GH51, which are inverting and retaining enzymes, respectively. The GH51 family is able to cleave Ara*f* bound both at the *O*-2 and the *O*-3 position of single substituted xylose residues. On the other hand GH43 can release a Ara*f* residue at the *O*-3 position of a doubly substituted xylose residue.[50] Hence, these enzymes can work in synergy making more substrate available for e.g. xylanases.

1.4.3 α-glucuronidases

Another class of accessory enzymes are the α-glucuronidases, which liberate MeGlcA or GlcA attached to the xylose backbone with inversion thereby releasing the β-glucuronic acid. The first glucuronidase family discovered was GH67, which can cleave the glycosidic bond between *O*-2 MeGlcA and xylose at the non-reducing end of short oligoxylans.[51] A more recently found family is the GH115, which is able to release MeGlcA and GlcA both on the terminal xylose residue and internally on polymeric xylans.[52,53]

1.4.4 Ferulic acid esterases

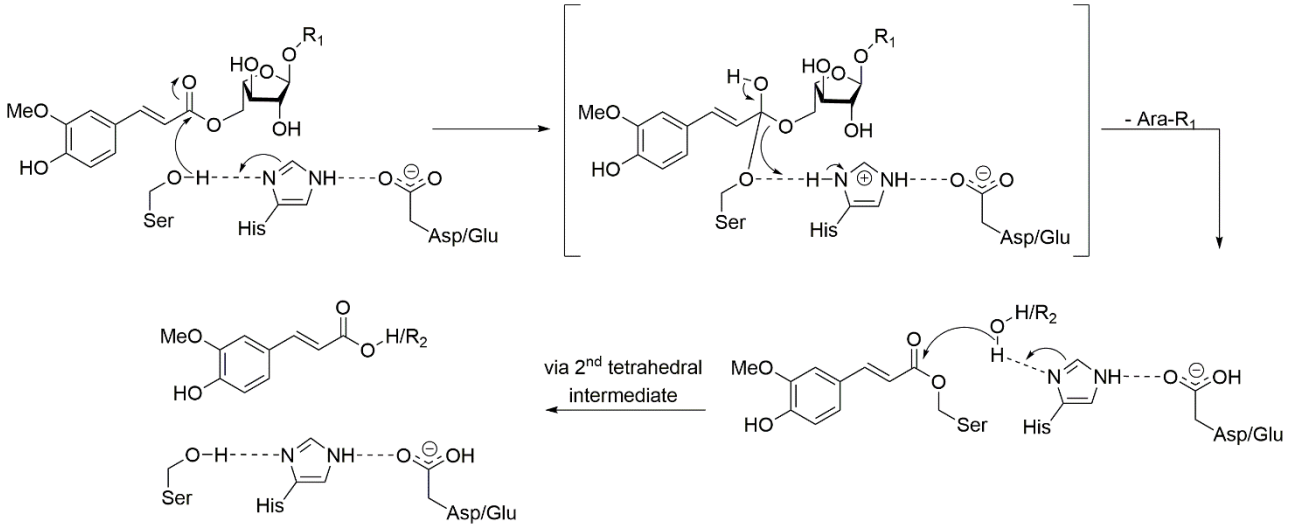
Ferulic acid esterases are a subclass of carboxylic acid esterases that are able to hydrolyse the ester linkage of a phenolic hydroxycinnamic acids and a poly- or oligosaccharide releasing the free hydroxycinnamic acid. Thus, FAEs have been developed in nature as accessory enzymes that aid other hydrolytic enzymes in biodegradation.[12,26,29,54] The first FAEs were discovered in the late 80'ies as esterases capable of releasing FA from xylan of e.g. wheat bran.[55–57] Since then FAE activity have been recognised as a common component of the hemicellulosic enzyme

system of many microorganisms, although the FAEs appear to be a very varied set of enzymes. Hence, similarity of gene sequences do not necessary guarantee the same function of the enzyme, which further complicates the classification.[58,59]

1.4.4.1 Reactivity of FAEs

The range for pH (pH 4 to 9) and temperature (20 – 65 °C) optimum of FAE activity is very broad, but most FAEs are mainly active at pH 4-7 and temperatures below 50 °C.[58,59] The FAEs contain a catalytic triad hydrolysing the substrate following the mechanism of serine proteases, where the triad consists of a conserved Ser-His-Asp/Glu structure.[60,61]

The catalytic mechanism of this triad involves the nucleophilic attack of the serine residue to form an acyl-enzyme via a tetrahedral intermediate. The nucleophilic attack is made possible through the acceptance of a hydrogen by the histidine residue stabilised by hydrogen bonding to the aspartic or glutamic acid residue of the triad. Following deacylation by nucleophilic water or another hydroxyl molecule from e.g. carbohydrates provide either the free hydrolysis product or a new ester through transesterification (Scheme 1-1).[61,62]



Scheme 1-1: Catalytic mechanism of the Ser, His, Asp/Glu triad of FAEs. R_1, R_2 = carbohydrate.

Even though the FAEs have the same catalytic triad there are a great many variations in the amino acid residues in close proximity to the triad and the substrate binding site thereby enabling differences in substrate specificity.[59,61]

1.4.4.2 Classification of FAEs

With little unifying sequence identity and structural characteristics FAEs appear to be a very diverse set of enzymes. Furthermore do esterases generally show a common trend of reacting

on a broad range of substrates. Several attempts have been made to try to categorise and classify these esterases. The first attempt was made by Crepin *et al.* in 2004.[63] This classification was based on activity profiles against four model methyl esters, the possible release of diFA, and the primary sequence identity of FAEs giving rise to a phylogenetic tree. This classification divided the FAEs into four sub-classes A, B, C, and D (Table 1-1).

Table 1-1: Classification of FAEs by Crepin *et al.*[63]

Sub-class:	Me *p*-coumarate	Me caffeate	Me ferulate	Me sinape	diFA
A	+	-	++	++	+
B	++	++	+	-	-
C	+	+	+	+	-
D	+	+	+	+	+

This classification is still the most commonly used in the literature by many researchers. However, as more FAEs have been discovered and their sequence elucidated this classification does not provide the complete picture. In 2008 de Vries and co-workers published a refined classification into seven subfamilies (SF) based on a phylogenetic analysis of the available fungal genomes at that time.[64] The classification showed that the FAEs does not have a common ancestor although they all demonstrate the same catalytic triad, but rather have evolved from highly divergent esterase families such as acetyl xylan esterases, tannases and lipases.[64] Then in 2011 Udatha *et al.* published a descriptor-based computational analysis with a pharmacophore model.[65] This study divided the FAEs into 12 families based on the whole amino acid sequence. Further inspection of the catalytic residue constellation of the sequences divided most of the found families into subfamilies (up to 4, A-D), which corresponded with previous specification regarding activity towards synthetic substrates of characterised FAEs of the individual family. The use of the characterised FAEs made the development of pharmacophore models of some of the subfamilies possible at the same time.[65] In 2016 de Vries and co-workers further developed their 2008 phylogenetic analysis.[27] A new phylogenetic tree was constructed using the amino acid sequence of 20 characterised FAEs and a genome mining strategy of 247 published fungal genomes as published in early 2015.[27] This search resulted in more than 1000 putative FAE sequences, which the phylogenetic analysis classified into 13 SFs (Figure 1-11). From the sequences, it can be seen that SF1-4 and SF9-11 are related to tannases. SF5 and SF6 show a relationship to acetyl xylan esterases, while SF7 is associated to lipases. A relationship to both lipases and choline esterases are seen for SF12 and SF13. From the

phylogenetic tree, it can be deduced that the substrate specificity of the FAEs does not reflect their evolutionary relationship. It was also suggested that the FAEs might be even further divided to support the ABCD classification when more FAEs have been characterised.[27]

To test the ability of the phylogenetic analysis to correctly divide the FAEs into SFs, 33 FAE sequences covering 11 SFs were selected for biochemical characterisation, including enzymes from five previously uncharacterised SFs (SF3, 8, 9, 10, and 12). Of the 33 sequences, six had previously been characterised and 27 were putative, it was found that 20 of these enzymes were active towards one or more of the four commonly used synthetic model substrates. The results displayed that members of SF1, 2, 5, 6, 7, 9, 10 and 13 process FAE activity towards methyl ferulate in varying degrees. Furthermore, it was found that SF9 members showed generally tannase activity. Activity towards methyl caffeate and methyl *p*-coumarate was observed for the tested members of SF3 and 12, respectively, but no activity towards methyl ferulate was found.[66]

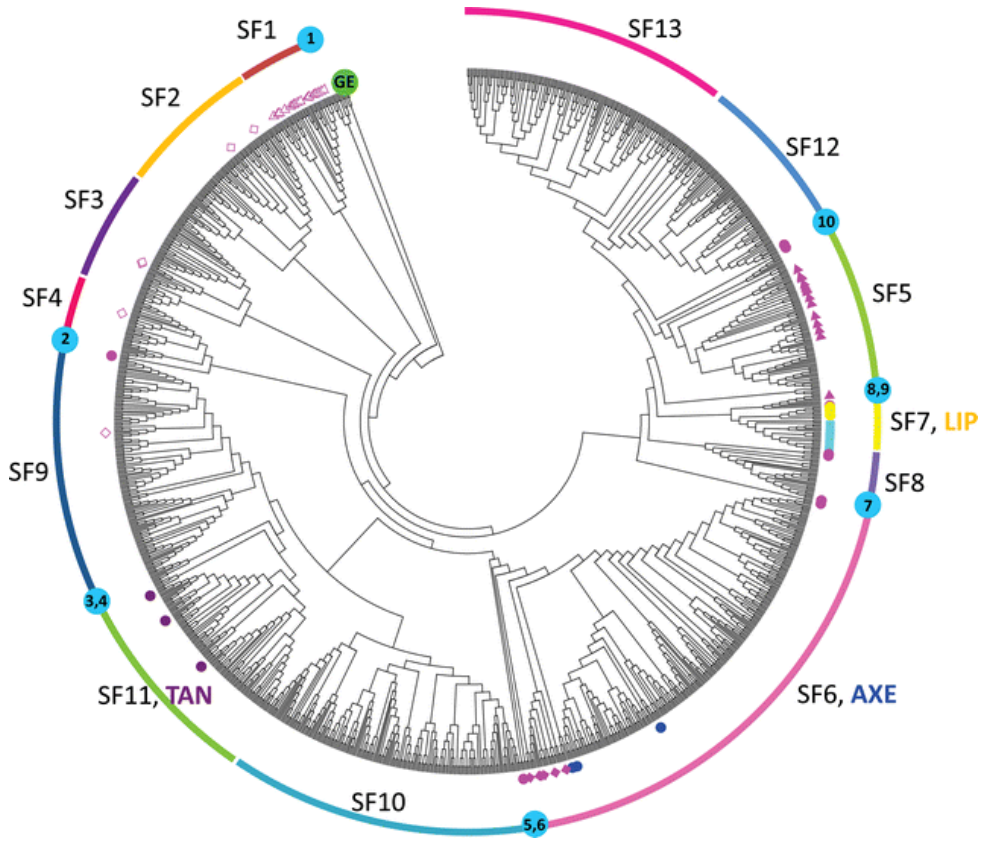


Figure 1-11: Phylogenetic tree as depicted by de Vries and co-workers.[27]

1.4.4.3 Applications of FAE

The FAEs have many applications in e.g. biofuel production and agriculture by co-operative enzyme activity in biomass degradation. Hence, the FAEs are already important tools, but

further investigation can lead to a better understanding of the many differences in specificity and reactivity, and consequently classification. As a further gain the released hydroxycinnamic acids in themselves also have application in food processing as flavour precursors and antioxidants in cosmetics as mentioned before.[58,67]

1.5 Chemical synthesis of xylan oligosaccharides

The continued study of the specificity, activity, and other biochemical properties of lignocellulosic enzymes will be much aided by the availability of well-defined target structures as written above. Current methods for obtaining such fragments often employ isolated polysaccharides obtained from natural sources by suitable pre-treatment procedures. Unfortunately the obtained polysaccharides are often inhomogeneous and analysis of the results from enzyme degradations become complicated especially with the complexity of highly substituted substrates.[68,69]

On the other hand, chemical synthesis of relevant oligosaccharide fragments allow for the preparation of specified targets with high purity and often in higher quantity, and consequently these can prove very valuable to modern research. The chemical community have already contributed with the synthesis of many oligosaccharide fragments. Challenges related to the synthesis of the fragments are usually associated with the stereospecific formation of the glycosidic linkages, the presence of branching units along the backbone, and protection group manipulation of the used building blocks and the target molecules. Hence, the synthesis of these targets often results in long and cumbersome synthetic pathways.

Synthesised fragments derived from xylan will be the main focus of this section. The published synthetic xylan fragments are summarised in Table 1-2, and the major efforts of the syntheses will be discussed in the following.

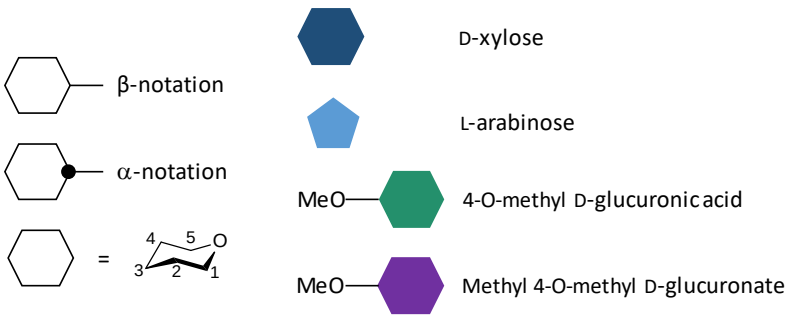
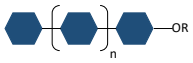
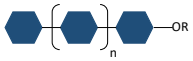
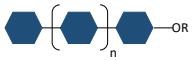
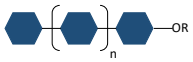
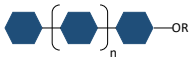
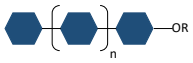
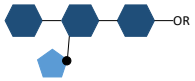
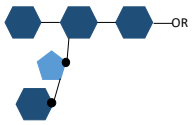
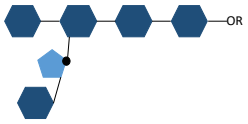


Figure 1-12: Glycan notation for Table 1-2.

Table 1-2: Schematic overview of published synthetic xylan fragments.

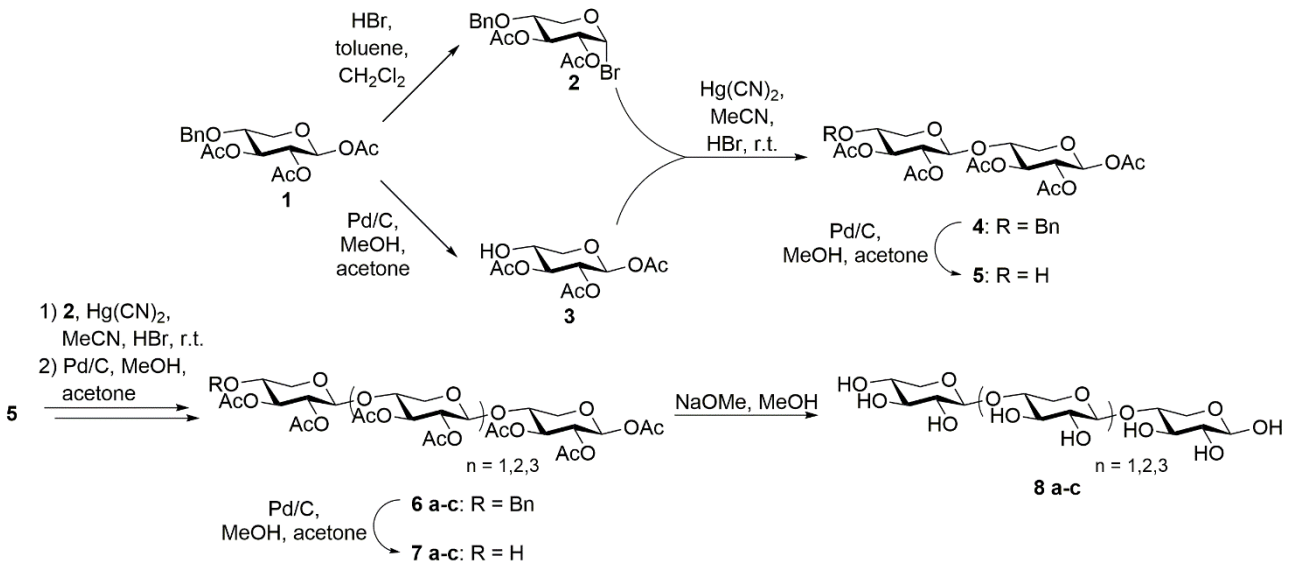
Entry	Structure		R	Year	Reference
1			H	1982	Hirsch *et al.*[70]
			Me	1982	Kováč *et al.*[71]
2		n = 2	H	1982	Hirsch *et al.*[70]
			Me	1982	Kováč *et al.*[71]
			H	1995	Takeo *et al.*[72]
			$(CH_2)_5NH_2$	2015	Schmidt *et al.*[73]
3		n = 3	H	1982	Hirsch *et al.*[70]
			Me	1982	Kováč *et al.*[71]
			H	1995	Takeo *et al.*[72]
4		n = 4	Me	1982	Kováč *et al.*[71]
			H	1995	Takeo *et al.*[72]
			$(CH_2)_5NH_2$	2015	Schmidt *et al.*[73]
5		n = 5	H	1995	Takeo *et al.*[72]
6		n = 6	H	1995	Takeo *et al.*[72]
			$(CH_2)_5NH_2$	2015	Schmidt *et al.*[73]
7		n = 7	H	1995	Takeo *et al.*[72]
8		n = 8	H	1995	Takeo *et al.*[72]
9			Me	1984	Hirsch *et al.*[74]
10			Me	1984	Hirsch *et al.*[74]
11			Me	1977	Kováč[75]
12			H	2001	Oscarson *et al.*[76]

13		$(CH_2)_5NH_2$	2015	Schmidt *et al.*[73]
14		$(CH_2)_5NH_2$	2017	Senf *et al.*[77]
15		$(CH_2)_5NH_2$	2015	Schmidt *et al.*[73]
16		$(CH_2)_5NH_2$	2017	Senf *et al.*[77]
17		Me	1980	Kováč *et al.*[78]
18		$(CH_2)_5NH_2$	2015	Schmidt *et al.*[73]
19		$(CH_2)_5NH_2$	2017	Senf *et al.*[77]
20		$(CH_2)_5NH_2$	2017	Senf *et al.*[77]
21		$(CH_2)_5NH_2$	2015	Schmidt *et al.*[73]
22		$(CH_2)_5NH_2$	2015	Schmidt *et al.*[73]
23		$(CH_2)_5NH_2$	2015	Schmidt *et al.*[73]
24		$(CH_2)_5NH_2$	2017	Senf *et al.*[77]
25		$(CH_2)_5NH_2$	2017	Senf *et al.*[77]
26		$(CH_2)_5NH_2$	2017	Senf *et al.*[77]
27		$(CH_2)_5NH_2$	2015	Schmidt *et al.*[73]

1.5.1 Synthesis of homoxylans

Several syntheses of xylan fragments using different synthetic approaches and glycosylation methods have been described over the last decades. The first reported synthesis of xylobiose was by Myhre and Smith in 1961.[79] The disaccharide was formed via a Helferich condensation[80] between benzyl 2,3-di-*O*-benzyl-D-xylopyranoside and peracetylated xylosyl bromide promoted by $Hg(CN)_2$. The Helferich method being a further development of the Koenigs-Knorr glycosylation, which was first published in 1901.[81] Both the Koenigs-Knorr and the Helferich methods originally featured the activation of glycosyl bromides either by silver or mercury salts, respectively. Since the first publications these methods with or without variations have become very popular and they are still widely used today. The popularity is being associated with the versatility that allows for the formation of 1,2-*trans* glycosides through neighbouring group participation.

Twenty years later a series of oligoxylans up to the pentaxylan was synthesised in a same fashion by Hirsch *et al.* (Table 1-2 entry 1-3).[70,82] 1,2,3-Tri-*O*-acetyl-4-*O*-benzyl-β-D-xylopyranose (**1**) was utilised as the common building block, which could be activated as a donor (**2**) by reaction with HBr or function as an acceptor (**3**) by selective deprotonation of the 4-*O*-benzyl group (Scheme 1-2). Mercury cyanide mediated the coupling between the two to afford the disaccharide as a 1:1.5 α/β mixture which could be separated by column chromatography to afford the desired β-product **4** in 35% yield. Subsequently the 4'-OH was liberated by hydrogenation to provide a new disaccharide acceptor (**5**). Repetition of the two last steps made the synthesis of tri-, tetra-, and pentaxylans (**8 a-c**) possible after saponification. The elongation of the oligoxylan did not affect the stereoselectivity and yields between 51% and 59% for the glycosylations were obtained. The same synthetic approach had furthermore previously been applied to obtain a series of methyl β-xylosides up to a xylohexoside (Table 1-2 entry 1-4).[71] In both cases it is seen that following this approach full stereoselectivity is not obtained even with the possibility of neighbouring group participation, and another drawback is the moderate yields obtained in the glycosylations.



Scheme 1-2: Synthetic strategy for the synthesis of homoxylans applied by Hirsch *et al.*[70,82]

A blockwise approach to the synthesis of a series of homoxylans up to xylodecaose were published by Takeo *et al.* in 1995 (Table 1-2 entry 2-8).[72] The synthesis utilised suitably protected thioxylosides as the donors which were coupled using *N*-iodosuccinimide (NIS) and silver triflate (AgOTf) as the promoting agents. This is an example of the use of thiodonors, which represent a popular choice in glycosylation chemistry, and they have had a leading role since they were first documented by Ferrier *et al.* in 1973.[83] Their success are being related to the ease of preparation, their stability, and easily controlled glycosylation protocols. The most common protocol for the stereoselective synthesis of 1,2-*trans* thioglycosides is the reaction of a peracetylated saccharide and the thiol of interest mediated by a Lewis acid, of which $BF_3{\cdot}OEt_2$ is the most commonly used, but several others are known.[84] Ferrier *et al.* originally employed phenyl thioglycosides as donors and mercury salts as the activating agents. Nowadays heavy metals are not much used, and many other activation options are used instead. These are often also more stereoselective, although at least a stoichiometric amount of the promotor is necessary. Some of the more popular choices today are iodide-based activating systems of NIS/TfOH or NIS/TMSOTf,[85] of which the Takeo synthesis[72] is an example. Additionally also thiophilic promotors have proved very successful. Examples of these are dimethyl(methylthio)sulfonium triflate (DMTST),[86] *p*-toluenesulfenyl chloride/AgOTf,[87] 1-benzenesulfinyl piperidine/triflic anhydride (BSP-Tf_2O),[88] diphenylsulfoxide/triflic anhydride (Ph_2SO-Tf_2O),[89] phenylsulfenyl chloride/AgOTf,[90] and the stable derivative *p*-NO_2PhSCl/AgOTf.[91] The advantages and disadvantages of the different promoting systems typically relate to the activation temperature of the donor and the stability of the promotor, e.g.

is NIS typically used for glycosylations around -30 °C where phenylsulfenyl chloride on the other hand activates the donors at -78 °C. Takeo *et al.* used the thiomethyl as the donor functionality for the synthesis of the xylan fragmens. The methyl 1-thio-β-xylobioside pentaacetate (**9**) and the methyl 1-thio-β-xylotrioside heptaacetate (**10**) acted as the glycosyl donors for construction of the terminal non-reducing ends (Figure 1-13).

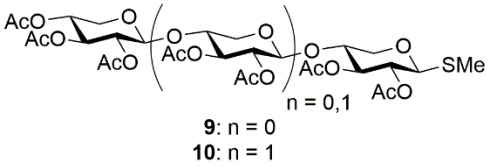
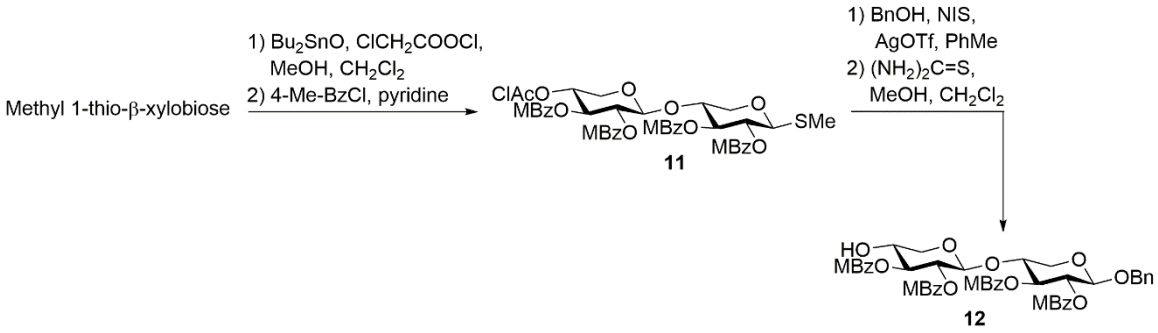


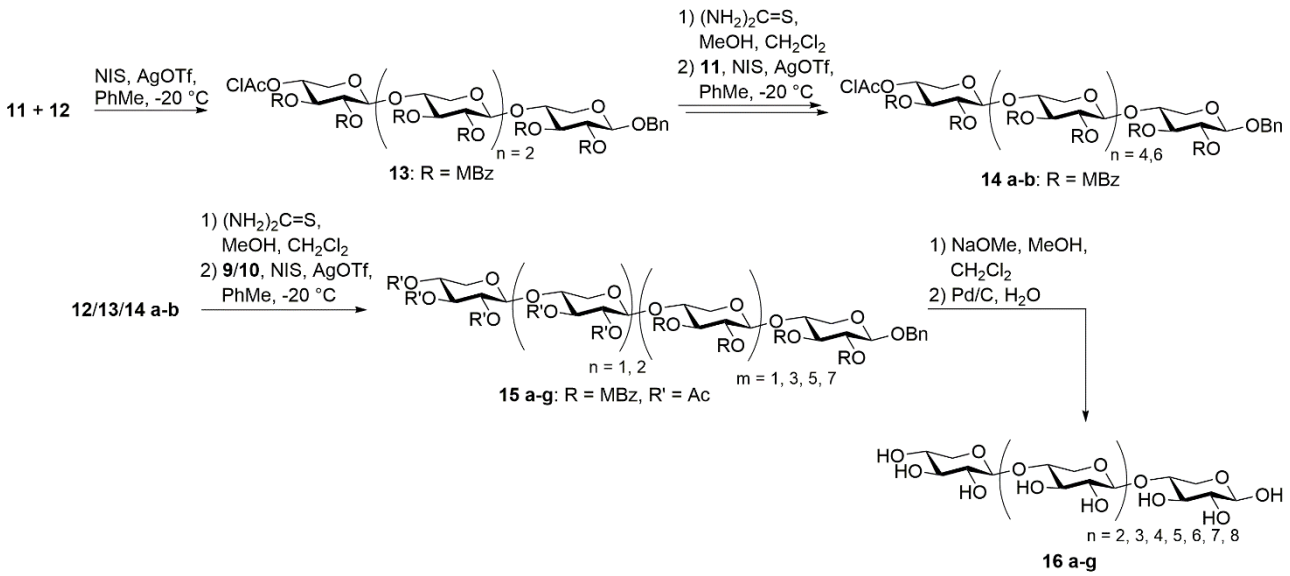
Figure 1-13: Donors employed by Takeo *et al.*[72]

Deacetylation of **9** provided the methyl 1-thio-β-xylobioside which could be converted into the thiodonor **11** in 70% through regioselective chloroacetylation followed by 4-methylbenzoylation (MBz) (Scheme 1-3). Condensation of **11** with benzyl alcohol followed by dechloroacetylation using thiourea provided the acceptor **12** for the terminal reducing end.



Scheme 1-3: Synthesis of disaccharide acceptors by Takeo et al. [72]

Building in the direction of the non-reducing end glycosylation of **12** with **11** promoted by NIS and AgOTf at -20 °C provided the tetrasaccharide acceptor **13**. Through repetition of dechloroacetylation and glycosylation, the hexa- and octasaccharide (**14 a-b**) were obtained (Scheme 1-4). Following dechloroacetylation and glycosylation of the acceptors **12**, **13**, **14 a** and **b** with the donors **9** or **10** under the same conditions as previously the tetra- to decaxylans (**16 a-g**) could be obtained after saponification and hydrogenolysis. Generally, the glycosylations were done in approximately 80% yield with full stereoselectivity for the β-linkage. Furthermore, the utilisation of the blockwise approach greatly limited the number of glycosylations and hence purifications and time of preparation.

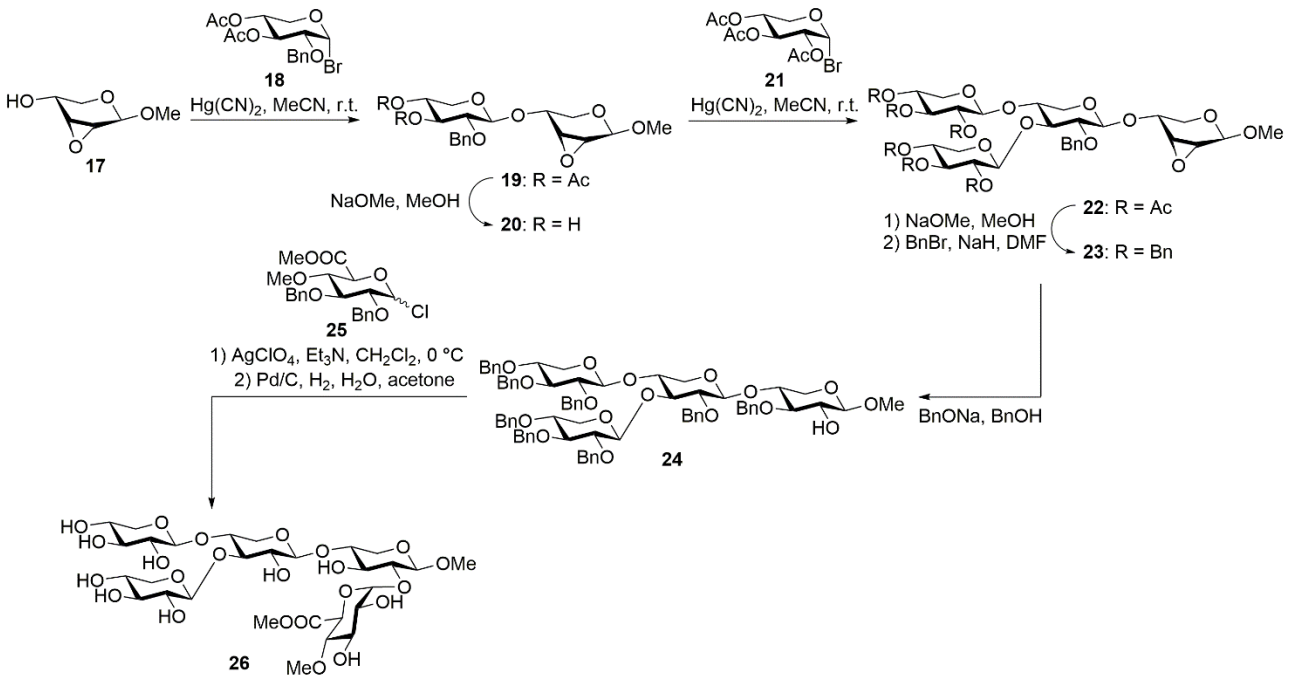


Scheme 1-4: Synthesis of homoxylans by Takeo *et al.*[72]

Thus, with the use of thiodonors greater stereoselectivity and yields were obtained. Furthermore the many possibilities for activation under different conditions and the ability of the thiodonors for interconversion provide a variety of alternatives via a common precursor.[92] Additionally, activation without addition of the acceptor, permits consecutive glycosylations in a one-pot fashion,[93] although none of these possibilities were employed by Takeo *et al.*

1.5.2 Synthesis of heteroxylans

In 1980 the stepwise synthesis of a xylotriose backbone with a β-(1→3)-linked xylose unit and a α-(1→2)-linked glucuronic acid was published by Kováč *et al.* (Table 1-2 entry 17).[78] The researchers started the synthesis by coupling of methyl 2,3-anhydro-β-D-ribopyranoside (**17**) and the xylosyl bromide donor **18** under Helferich conditions in 62% (Scheme 1-5). The resulting disaccharide **19** was then selectively deprotected at the *O*-3 and *O*-4 positions and further glycosylated with xylosyl bromide **21** affording the tetrasaccharide **22** in 40%.[94] After exchange of the acetyl protections groups for benzyls, the epoxide was selectively opened with benzyl alcohol according to the Fürst-Plattner rule[95] affording the tetrasaccharide **24** with *xylo*-configuration. Following the free *O*-2 hydroxyl was glycosylated with glucuronate donor **25** in the presence of silver perchlorate and 2,4,6-collidine. The α-adduct was formed as the main product in 57%. Complete deprotection yielded the pentasaccharide **26**.[78]



Scheme 1-5: Synthesis of the branched xylotetraose with a (1,2)-4-OMe-α-D-glucuronic acid.[78,94]

The same stepwise synthesis had previously been applied for synthesis of a xylobiose with GlcA attached at the 2-position (**27**) in 1977 (Table 1-2 entry 11),[75] being the first example of the synthesis of a heteroxylan oligosaccharide with a α-(1→2)-glucuronic acid (Figure 1-14). The trisaccharide was built from disaccharide **20**, which instead of glycosylation with **21** was directly benzylated, the epoxide regioselectively opened and then coupled with donor **25**. The glycosylation provided the α-linked **27** as the main product in 87% yield, showing a greater selectivity for the α-linkage (α/β 7.7:1) than in the later example. Glycosylations with GlcA will be further discussed *vide infra*.

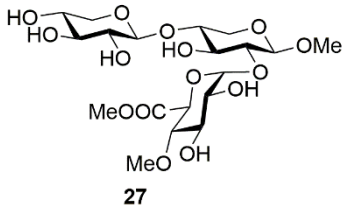
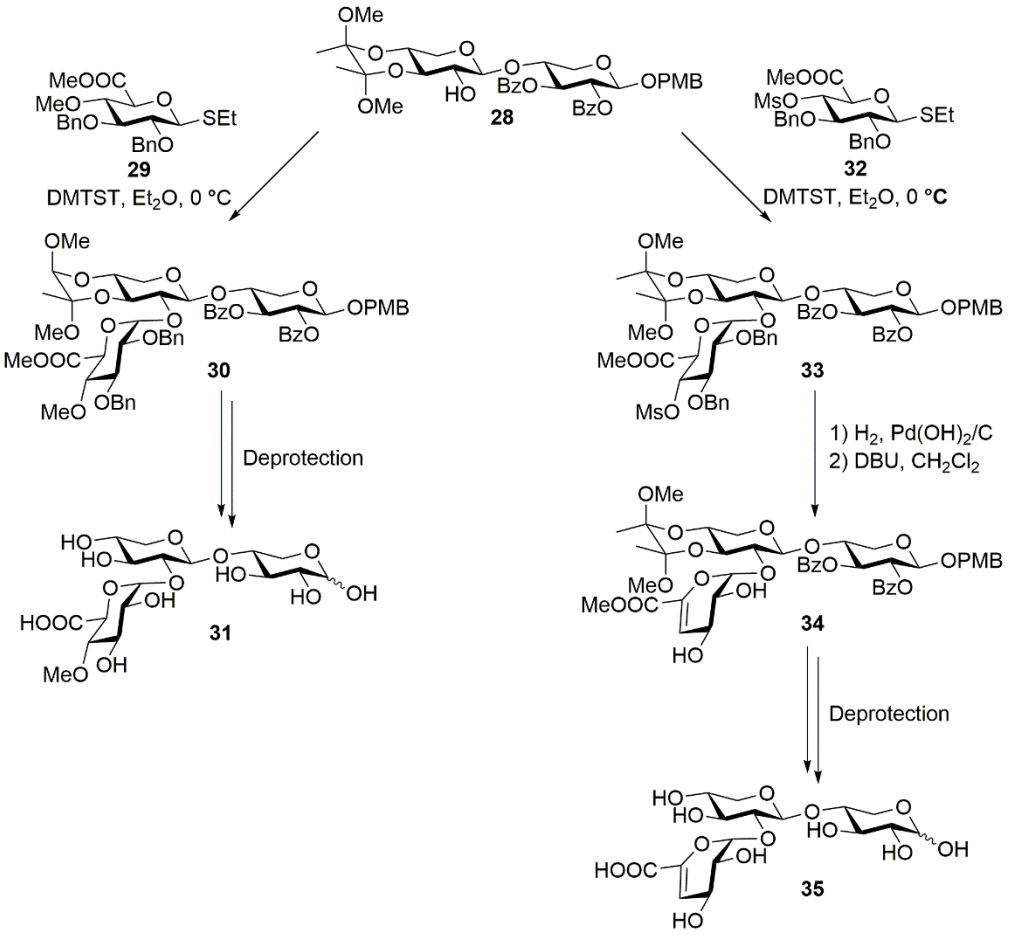


Figure 1-14: The first reported synthesised heteroxylan related to GX.[75]

Oscarson and Svahnberg reported another example of short glucuronoxylan fragments based on xylobiose in 2001 (Table 1-2 entry 12).[76] One of the trisaccharides had a α-(1→2)-linked glucuronic acid (**31**), the other having a α-(1→2)-hexenuronic acid (**35**), which can be formed under alkaline pre-treatment of wood for paper production as mentioned earlier. The common xylobioside acceptor **28** with a free 2'-OH was coupled with the two glucuronic acid donors,

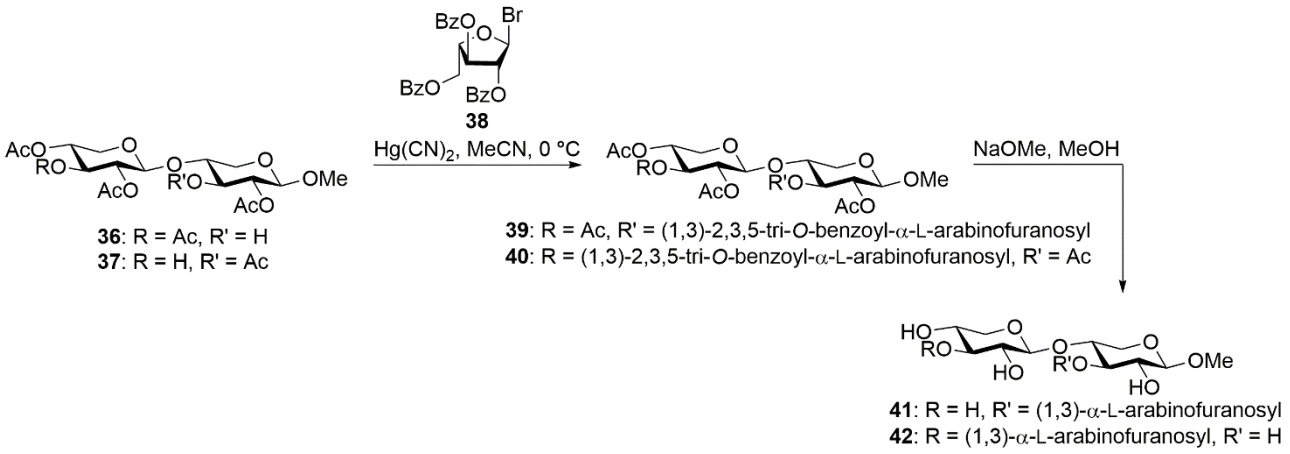
which were differentiated at the 4-position by having either a methoxy (**29**) or a mesyl (**32**) group. The glycosylations were promoted by DMTST in diethyl ether, which afforded the α-linked trisaccharides **30** and **33** as the only products in 89 and 80% yield, respectively (Scheme 1-5). The fully deprotected product **31** was formed in 32% yield after protection group maninpulations and deprotections. The protected trisaccharide **33** underwent β-elimination after hydrogenolysis to form the α,β-unsaturated uronic derivative **34** in 72% yield. The following protection group manipulations and deprotections afforded the fully deprotected trisaccharide **35** in 45% yield (Scheme 1-6). This is the first examples where also the methyl ester of GlcA have been deprotected providing the actual carboxylic acid derivative.



Scheme 1-6: Synthesis of GX fragments by Oscarson and Svahnberg.[76]

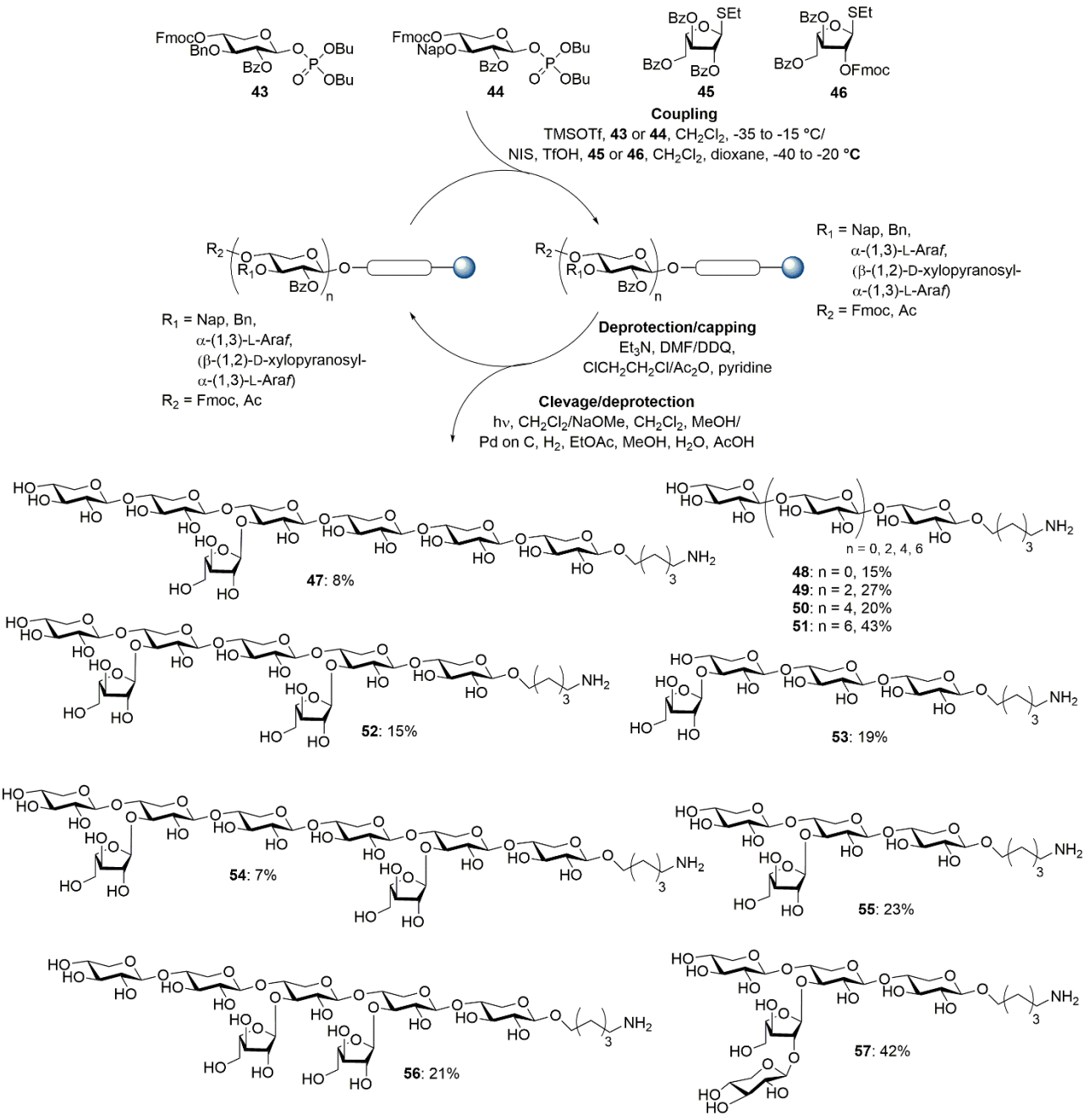
The first reported synthesis of heteroxylans related to AX was published in 1984 by Hirsch and Petráková (Table 1-2 entry 9 and 10).[74] These structures were based on xylobiose with an *O*-3-linked arabinofuranosyl moiety on either the reducing or the non-reducing end. The synthesis was based on glycosylation of two appropriately protected xylobiose fragments **36**[96] and **37**[97] with the arabinose bromide donor **38** promoted by mercury cyanide (Scheme 1-7). The

glycosylations were achieved in more than 90% yield with complete stereoselectivity. The deprotected trisaccharides **41** and **42** were obtained after saponification. This is the first example of xylan fragments synthesised using the Helferich method where complete stereoselectivity was achieved and this can possibly be explained by the donor being an arabinofuranosyl group as well as the possibility of neighbouring group participation.



Scheme 1-7: Synthesis of two trisaccharides related to arabinoxylan.[74]

Recently, Pfrengle and co-workers have developed an automated synthesis of arabinoxylans consisting of xylan oligosaccharides up to an octasaccharide with one to two α-(1→3)-linked arabinofuranosides in different positions of the backbone (Scheme 1-8) (Table 1-2 entry 2, 4, 6, 14, 15, 18, 21-23, 27).[73,77] Two different xyloside building blocks **43** and **44** which differentiated in the protection of their *O*-3 position by having the permanent benzyl group or the temporary 2-naphthylmethyl (Nap) group served as building blocks for the linear xylan backbone. For substitution, the perbenzoylated and 2-*O*-Fmoc-L-arabinofuranosides **45** and **46** were used. The temporary Fmoc group of **46** made further xylan substitution possible. The donor functionalities of the xylan building blocks were dibutylphosphates, which were activated with TMSOTf, while the arabinose building blocks featured thioethyl glycosides activated by NIS/TfOH. The chains were build towards the non-reducing end and the reducing end was linked to a functionalised resin. Cleavage of the linker after synthesis revealed the eleven AX fragments (**47**-**57**) as a conjugation tool in 7-43% yield.



Scheme 1-8: Automated solid-phase synthesis as presented by Pfrengle and co-workers in 2015 as well as the obtained structures.[73]

In 2017 the method was further optimised by the addition of the xylan building blocks **58** and **59** which made the synthesis of the AX fragments **60**-**66** possible in 4-41% yield (Figure 1-15) (Table 1-2 entry 13, 16, 19, 20, 24-26). The xylan donor **58** allowed for mono-α-(1→2)-substitution with arabinose through the 2-*O*-(azidomethyl)benzoyl (Azmb) temporary protection group. Using the donor **59**, which was orthogonally substituted with 2-*O*-Azmb and 3-*O*-Nap, the synthesis of disubstituted arabinoxylan **65** was possible. All the AX fragments

from both the 2015 and 2017 publications were printed on microarray slides and used for probing of anti-xylan monoclonal antibodies.

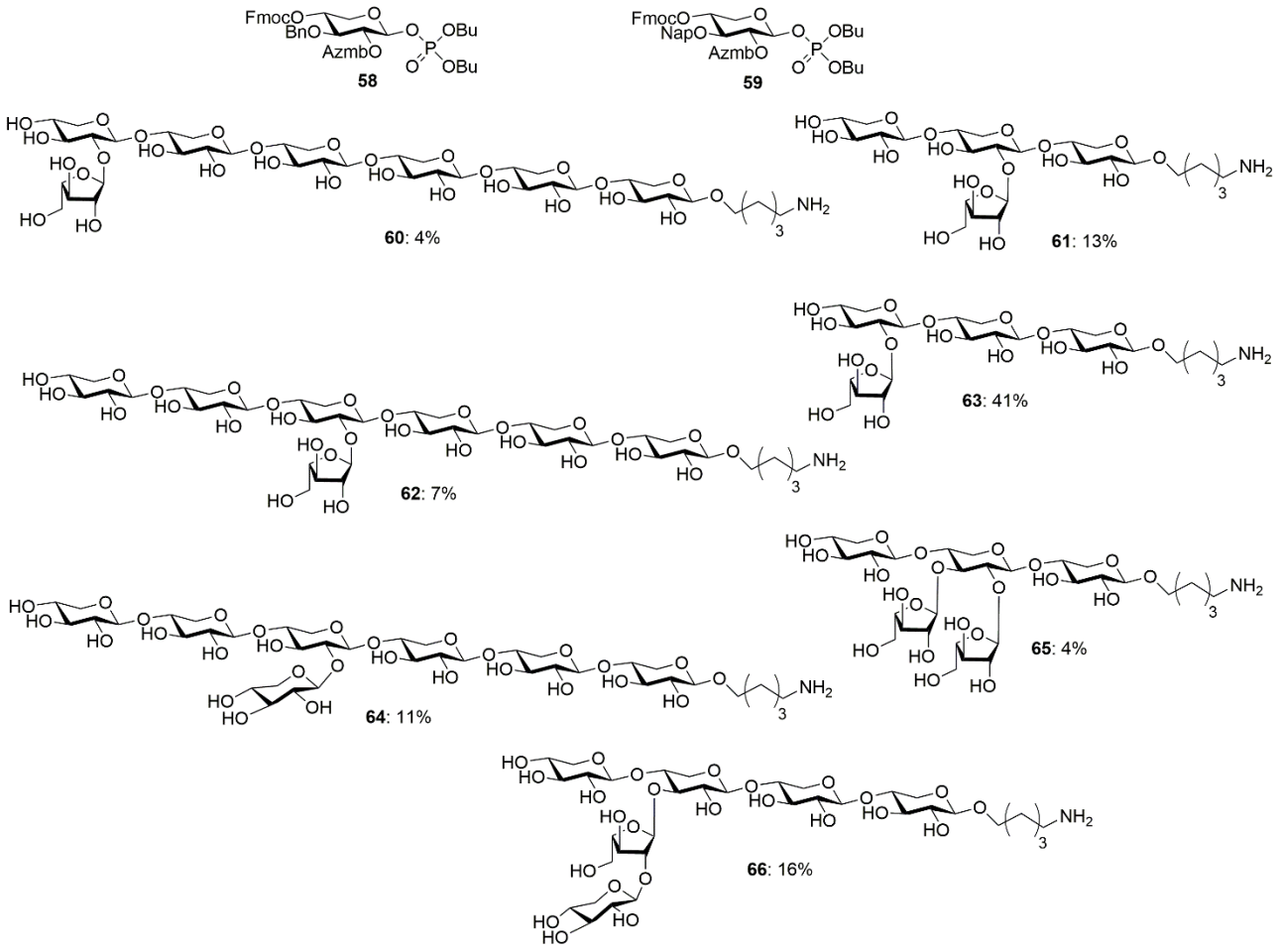


Figure 1-15: Additional xylose donors and AX fragments synthesised by Pfrengle and co-workers in 2017.[77]

The automated solid-phase synthesis of AX fragments represents a fast route to these structures, which yields many different structures, and furthermore is the purification easy due to the solid-phase protocol. The drawbacks nevertheless being the use of specialised equipment, extended optimisation of the method before sufficient practicality is achieved, and very low yields of some of the structures are still observed as well as generally small amounts of the targets are obtained (1-4.4 mg for all structures except for **66** of which 15.5 mg was obtained). However, the method represents a very interesting leap forward in the synthesis of AX fragments, which might be further developed.

1.6 Glycosylations with glucuronic acid residues

Polysaccharides with uronic acids are widespread in nature, and the ability to form these linkages synthetically is therefore very important. Generally two different approaches have

been used, one being post glycosylation oxidation, where the carboxylic acid moiety is formed after the critical glycosylation. Hence, this approach requires extra steps for forming the carboxylic acid and the need of an orthogonal protection group on the primary alcohol. The other approach involve pre-glycosylation oxidation, although a lower reactivity of the uronic acid both as a donor and an acceptor has to be accommodated, which is due to the electron-withdrawing effect of the ester. Furthermore, side reactions such as epimerisation of the *C*-5 and β-elimination have to be avoided.[98,99]

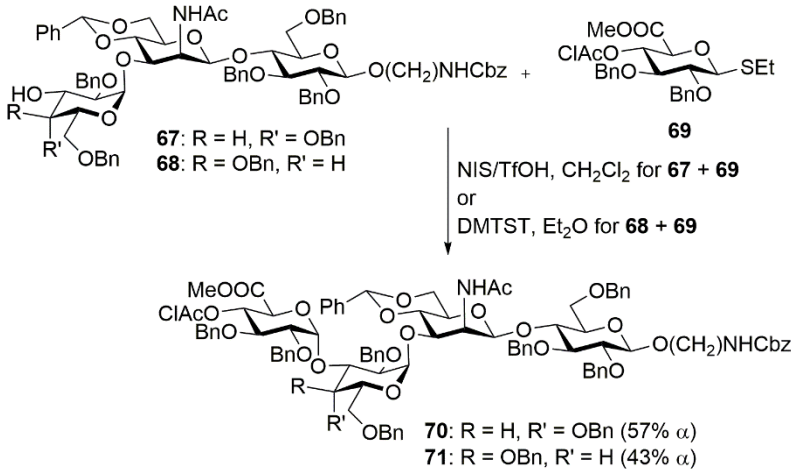
As shown above there are some examples in the literature regarding the synthesis of GX fragments (Section 1.5.2), where all are using a pre-glycosylation oxidation approach, but applying different donor functionalities and activation strategies for glycosylation with GlcA. The GlcA in GX and GAX is α-(1→2)-linked, and in synthesis the most common way of obtaining this type of 1,2-*cis* linkage is through the use of a non-participating protecting groups on *C*-2. In the cases shown, the benzyl group have been used, but the glycosylation often is still difficult. When the oxycarbenium ion is formed after activation of the anomeric leaving group and there is no participating protection group, then there are two stereoselective outcomes; an α- or a β-linkage. Under thermodynamic conditions, a general preference for the α-anomer is observed, which was first noticed by J. T. Edward.[100] Two alternative explanations have been given for the origin of this so-called anomeric effect. The first one is based on the favourable dipole-dipole interaction, which is present for the α-anomer. The other explanation is the stabilisation of the α-anomer by the n(O) → σ*(C-X) orbital interaction.[101] However, the idea of the anomeric effect as an explanation for α-selectivity under kinetically controlled glycosylations (which most glycosylations are) was refuted in 2012 by Ian Cumpstey.[102] He argued that it is the structure of the individual carbohydrate units, which determine the selectivity in glycosylations under kinetic control.

However, another very common method for influencing the stereoselective outcome of a glycosylation is the use of different solvents explained by the so-called solvent effect.[103] One of the advantages being that the solvent can be easily removed during work-up, whereby it is possible to avoid special protection groups, which furthermore simplifies the synthesis of the building blocks for glycosylations.

Apart from physical limitations arising from the melting and boiling points of the solvents, the polarity has a great influence on the solvent effect. The solvation of the ion pair after activation of the donor is poor in weakly polar solvents, and these solvents are regarded as non-

dissociating solvents. Et_2O and dioxane are weakly polar and non-dissociating solvents, where nitromethane and acetonitrile are highly polar and dissociating solvents. Dichloromethane on the other hand is only moderately polar. Furthermore, the electron-donating capacity, which refers to the interaction between the solvent and the solute, has an influence. The electron-pair-donation of the solvent towards the anomeric center of the oxycarbenium ion stabilises the ion pair, hence promoting α-glycosylation.[103] Et_2O is an example of a coordinating solvent, and since this solvent is also weakly polar it is a good solvent for α-glycosylations, which can explain the choice of this solvent in the Oscarson and co-workers protocol.[76] Dichloromethane, however, is non-coordinating but only moderately polar, which might be one of the reasons for the relative high α-stereoselectivity obtained by Kováč.[75,78]

As an example of the unpredictable GlcA-glycosylations, the synthesis by Oscarson and co-workers in 2001 showed full α-selectivity when using the thioethyl donor, a non-participating protection group on *C*-2, and activation by DMTST in Et_2O, whereas activation by NIS/TfOH in dichloromethane gave the β-product in 62% (Scheme 1-9).[76] Hence the system seemed to be very dependent on the promotor and the solvent. However, the completely reversed selectivity was later observed also by Oscarson and co-workers in 2003 when synthesising the repeating unit of *S. pneumonia* type 9A.[104] The GlcA donor **69** was coupled with either trisaccharides **67** or **68** (Scheme 1-8). Coupling of **67** with **69** using DMTST in Et_2O mainly gave the β-product, while the NIS/TfOH system in dichloromethane gave the best results in regard to stereoselectivity and yield (2:1 α/β). However, for coupling of **68** with **69** the best results was obtained by the DMTST promotor system although with lower selectivity (3:2 α/β).



Scheme 1-9: Stereoselectivity and reaction conditions employed by Oscarson and co-workers in 2003.[104]

Hence, even when using the same promoting system, solvent, and donor functionality the glycosylation with GlcA can be difficult to predict.

2. Chemical synthesis of arabinoxylans fragments

This chapter covers the results obtained for the synthesis of arabinoxylans fragments, where both former and own results will be described and discussed. The project was initiated by a former Ph.D. student in the group Maximilian Felix Böhm, who performed the optimisation of the glycosylation protocol.[105] Later we worked in parallel with the goal of achieving different target structures. The first part of this chapter will illustrate the optimisation of the glycosylation protocol done by M. F. Böhm during his Ph.D. study.[105] The second part will review the synthesis of three arabinoxylans target structures produced for this thesis.

2.1 Aim of the project

Based on the continuous need for pure and well-defined oligosaccharide substrates for elucidation of hemicellulosic enzymes it was decided to pursue the synthesis of three AX fragments (Figure 2-1). These fragments consisted of a common β-(1→4)-linked xylose backbone differing in the substitution pattern. The synthesis was designed in a way that made it possible for the xylan backbone to be single substituted with an α-(1→3)-linked Ara*f* or double substituted with an α-(1→2)- and an α-(1→3)-linked Ara*f* residues through the use of orthogonal protection groups. The synthetic strategy should be reliable and allow for the

construction of the AX fragments but through minor modification also allow for the synthesis of e.g. GX fragments due to the common backbone.

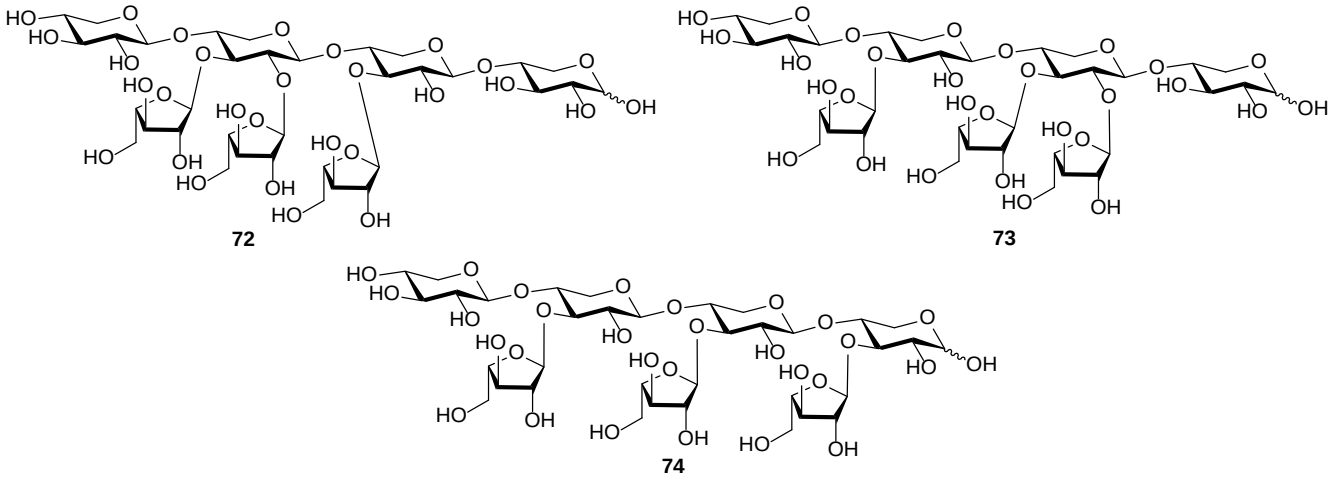


Figure 2-1: Structures of the three chosen target structures for this project.

The specific structures were chosen since the differing in substitution pattern can help validate the significance of the different patterns for enzyme specificity of both xylanases and arabinofuranosidases or assign the type of glycosyl hydrolase for novel enzymes (Figure 2-2). As an example would incubation of e.g. target structure **72** with a GH10 or GH11 not lead to enzymatic breakdown unless the target previously had been subjected to cleavage with GH43 and GH51. However, enzymes belonging to GH5 would be able to cleave the target directly, hence different types of xylanases could be distinguished based on ability to act on the target. It would also be possible to distinguish between α-arabinofuranosides of GH43 and GH51, depending on which Ara*f* unit would be cleaved off. For comparison between targets, a GH51 would be able to cleave of all three Ara*f* units on target structure **74**, but not the on target structure **72**.

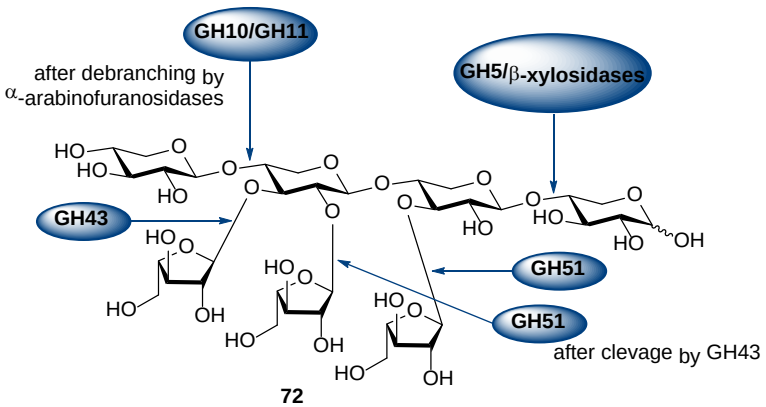
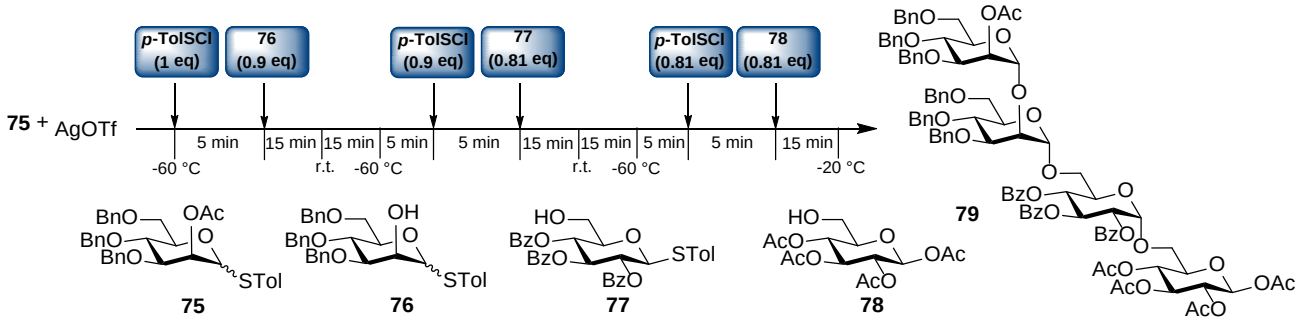


Figure 2-2: Example of possible enzymatic degradation of target structure **72** with glycosyl hydrolases.

2.2 Optimisation of the glycosylation protocol

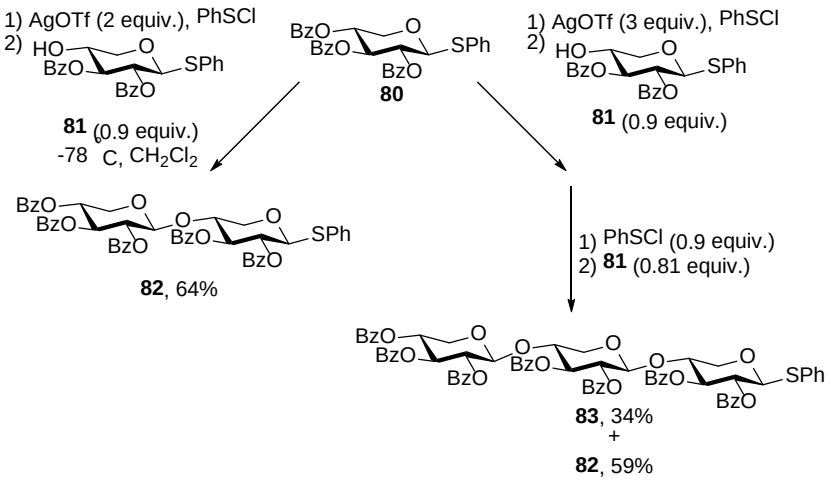
The research reviewed in this section (section 2.2) represents work done by Maximilian Felix Böhm.[105]

The first approach attempted for the glycosylation was based on the work of Huang *et al.*[87] In 2004 they described an optimised one-pot strategy using thioglycosides that could be preactivated with *p*-tolylsulfenyl chloride and silver triflate before reaction with another thioglycoside. Afterwards the resulting thioglycoside could be activated and a new glycosylation round could be performed. Hence, it was possible to construct the tetrasaccharide **78** in 55% yield in a fairly short time (Scheme 2-1).



Scheme 2-1: One-pot approach as published by Huang *et.al.*[87] eq: equivalents.

The same conditions as described by Huang[87] was applied for the synthesis of the disaccharide **82** in 64% yield using the building blocks **80** and **81** with benzoyl group as permanent protection groups (synthesis of the compounds and choice of protection group will be discussed *vide infra*) (Scheme 2-2). However, applying the conditions as a one-pot protocol for the synthesis of trisaccharide **83**, the desired trisaccharide was obtained in only 34% yield while the major product was the disaccharide **81** (59%) (Scheme 2-2).

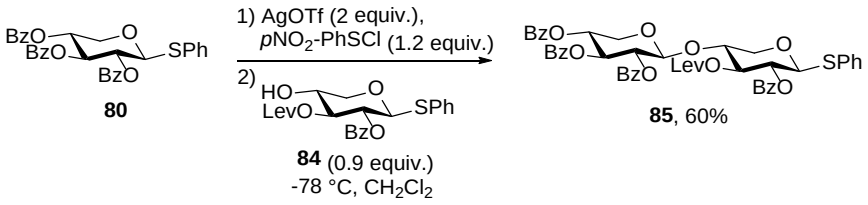


Scheme 2-2: Formation of disaccharide **82** as well as the one-pot synthesis of trisaccharide **83**[105] following the protocol as described by Huang.[87]

Then M. F. Böhm spend efforts on improving the yield by screening of different promotor systems: Me_2S_2/Tf_2O,[106] 1-benzenesulfinyl piperidine/Tf_2O,[88] and *p*-nitrobenzenesulfenyl chloride/AgOTf.[91] Only the last system led to a clean conversion as well as disaccharide **82** in 79% yield, and during a one-pot protocol using a slight excess of the promotor (1.2 equiv.), trisaccharide **83** could be isolated in 69% yield (results not shown).

The strategy was further developed to include an acceptor containing two different protection groups allowing for later attachment of an arabinose unit at the *O*-3 position. Several temporary protection groups were considered. The first candidate being considered was benzyl. According to a procedure by Huang *et al.*[107] persilylated *p*-methylphenyl 1-thio-xyloside can be selectively benzylated on the 3-position followed by silyl deprotection. Subsequently an ytterbium triflate catalysed benzoylation was reported to be selective at the 2-position. Unfortunately, these results could not be reproduced by M. F. Böhm and only traces of the 3-benzylated product could be isolated (results not shown).[105]

The second temporary protection group candidate, which was considered, was an *O*-3 chloroacetyl group, and the permanent protection group benzoyl at *O*-2 (introduction of the *O*-2 group will be discussed *vide infra*). However, when this acceptor was used under the previously established glycosylation conditions the corresponding disaccharide was obtained in only 42% yield (results not shown).[105] The main reason for this low yield was ascribed to the lower stability of the chloroacetyl group in the presence of the silver salt. Lastly, the levulinyl (Lev) group was chosen as the temporary group at *O*-3, as glycosylation with **80** and the acceptor **84** produced the disaccharide **85** in 60% yield (Scheme 2-3).[105]



Scheme 2-3: Glycosylation with acceptor **84** with a Lev group in the 2-position as done by M. F. Böhm.[105]

However, while M. F. Böhm worked with the optimisations of the glycosylations a sudden drop in yield of about 20% was observed. A colour change of the commercially available promoter *p*-NO_2PhSCl prompted the suspicion of a change in purity. Repeated attempts to recrystallise the compound failed. Sublimation, however, worked very well, after which a bright yellow compound could be obtained.[105] The sublimated product can be stored in the freezer under an atmosphere of nitrogen for two month obtaining comparable yields in subsequent glycosylations. With the sublimated promotor, it is furthermore possible to use exactly one equivalent for the preactivation instead of 1.2 equivalent of the commercially available material making sure none of the acceptor is activated during the glycosylation. However, the purification of the promotor did not make a change with regard to the yield, but allowed a constantly good quality of the promotor for glycosylations.

Excluding impurities and moisture in the reaction as a cause for the drop in yield the only variable left to be studied was the temperature. To allow for exact temperature measurements during the glycosylation a 2-neck Schlenk flask was utilised for the investigation of the temperature influence. The temperature influence was tested during the glycosylations forming disaccharides **82** and **85**. It was found that the activation of the donor is quick and goes to completion within 5-10 min even at -78 °C. Furthermore, the investigations showed that when adding the acceptor at temperatures lower than -60 °C, aglycon transfer becomes increasingly dominant leading to the re-generation of donor **80** (aglycone transfer will be discussed *vide infra*). This was concluded by observations via thin layer chromatography (TLC) and reisolation of **80** in approximately 10% yield after column chromatography, alongside isolation of the disaccharides **82** and **85** in 54% and 44%, respectively (results not shown).[105] M. F. Böhm concluded that while it is advantageous to keep the temperature as low as possible during the activation to ensure the stability of the intermediate triflate,[88] it is necessary to raise the temperature to more than -60 °C when adding the acceptor to ensure the desired reaction pathway. After several trials to examine the temperature influence, the optimum reaction temperature was found to be around -55 °C for the first 15 minutes (Figure 2-3).[105]

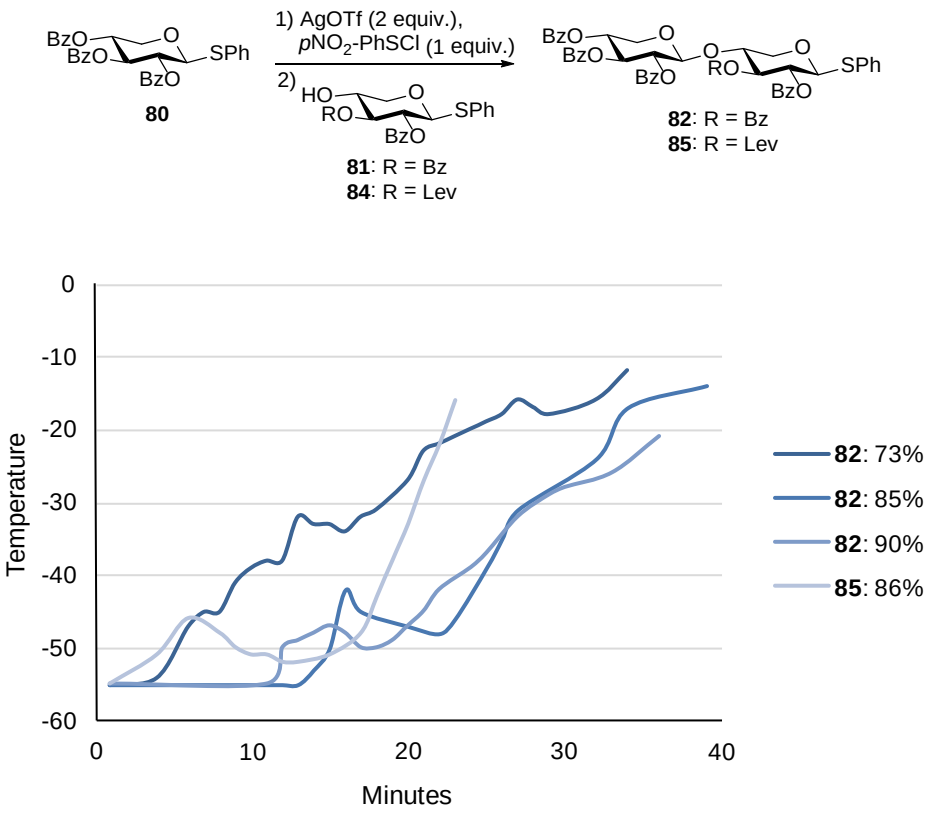
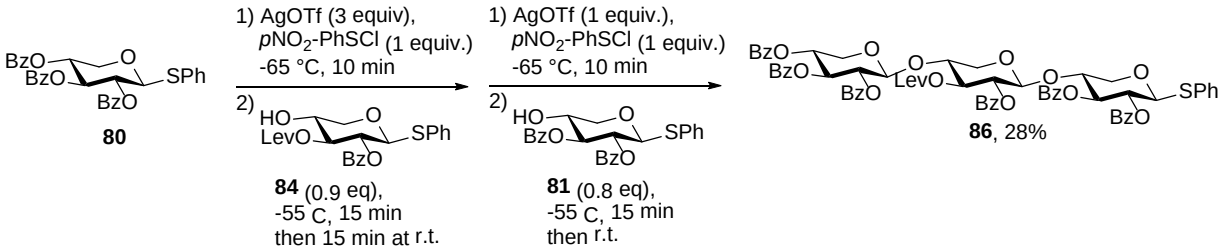


Figure 2-3: Influence of the reaction temperature on the yield of glycosylation providing disaccharides **82** or **85** as reported by M. F. Böhm.[105]

When M. F. Böhm applied the optimised reaction conditions (promotor and temperature) as a one-pot approach using acceptors **81** and **84** to obtain trisaccharide **86**, only a 28% yield of **86** was reached (Scheme 2-4). Running the same reaction with addition of 2,4,6-tri-*tert*-butylpyrimidine to diminish the possible influence of the increasing amount of acid,[108] lowered the yield even more to 17%, while 80% of the acceptor **81** was isolated (results not shown). Furthermore, TLC showed a complex mixture.[105]



Scheme 2-4: One-pot procedure applying the optimised reaction conditions.[105]

As the one-pot strategy with pre-activation did not seem viable without a major time investment, a consecutive approach also depending on pre-activation was investigated by M. F.

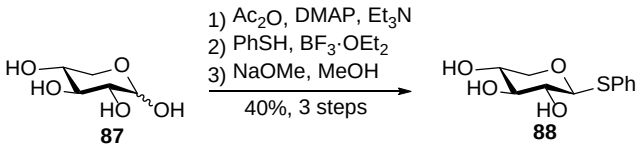
Böhm. Disaccharide **84** was pre-activated and reacted with **80** providing trisaccharide **85** in 69% yield. Although this approach is not as fast as the one-pot procedure, the glycosylation products once isolated can be immediately used as donors in the next glycosylation step without any further modification, thus leading to better yields and still saving time compared to approaches not depending on the pre-activation principle.[105]

2.3 Synthesis of monomeric xylose building blocks

Research presented in this and the following sections of this chapter represent work done by the author for this thesis.

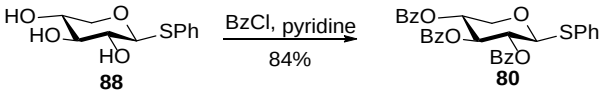
Some of the monomeric xylose building block have already been introduced (*vide supra*), but the synthesis of each is being presented in detail here. Due to the common β-(1→4)-linked xylose backbone the building blocks should have a selectively removable moiety at the *O*-4 position (a *para*-methoxybenzyl group). The protection group at *O*-2 should be able to act as a participating neighbouring group (an ester). Furthermore, the protection group at *O*-2 and *O*-3 should differ for the building blocks having either a permanent protection group (benzoyl group) or a temporary group (levulinyl group) as described in section 2.2, which make selective removal possible for later installation of arabinose building blocks. Moreover, the thiophenyl group was chosen as the common donor functionality due to the great stability and the ease of activation, which made the optimised glycosylation protocol possible.

The triol **88** represents a common starting material for the synthesis of all the monomers needed for the synthesis of the selected targets. Compound **88** can be prepared in three steps from commercially available D-xylose in easy and scalable procedures (Scheme 2-1). The first step is a standard peracetylation protocol in the presence of the catalyst *N,N*-dimethyl-4-aminopyridine (DMAP), following a thio-glycosylation promoted by $BF_3{\cdot}OEt_2$ and a final deprotection using Zemplén conditions provides compound **88**.[109,110]



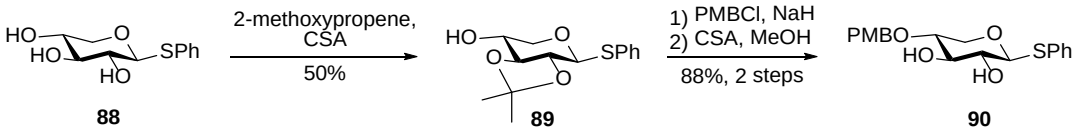
Scheme 2-5: Synthesis of **88** from D-xylose.

Perbenzoylation[111] of **88** afforded the donor **80**, which is to be used as the non-reducing end building block (Scheme 2-6).



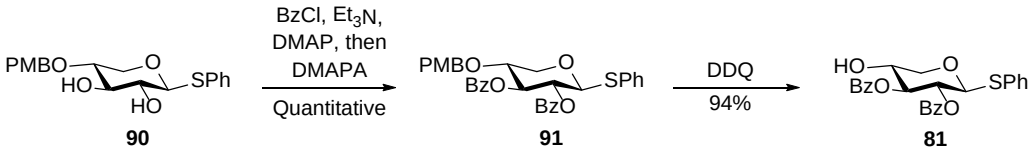
Scheme 2-6: Formation of building block **80**.

For the other monomers, further protection group manipulations were necessary to obtain bifunctional building blocks. As written above a PMB ether was chosen as a removable moiety for the *O*-4 position. Specifically the PMB ether was chosen as it can be selectively removed in the presence of 2,3-dichloro-5,6-dicyanobenzoquinone (DDQ).[112] However, the hydroxyl groups of xylose have similar reactivity, even though the general reactivity of β-xylosides is recognised as *O*-4>*O*-3>*O*-2.[113] Thus, to be absolutely sure about the regioselectivity in the introduction of the PMB functionality at the *O*-4 position, the *O*-2 and *O*-3 positions were protected as the 2,3-isopropylidene. This was done by reaction of **88** with 2-methoxypropene in the presence of camphorsulfonic acid (CSA) (Scheme 2-7).[114] The reaction afforded the xyloside **89** (50%) as the major product, although the 3,4-acetonide and the fully protected species formed by one further acetal at the *O*-4 position of **89** were also formed. This rendered the purification challenging, as the regioisomers are difficult to separate. Subsequently the *O*-4 position of **89** is protected as the PMB ether under basic conditions, followed by acidic hydrolysis of the acetal leading to the diol **90** (Scheme 2-7).[115]



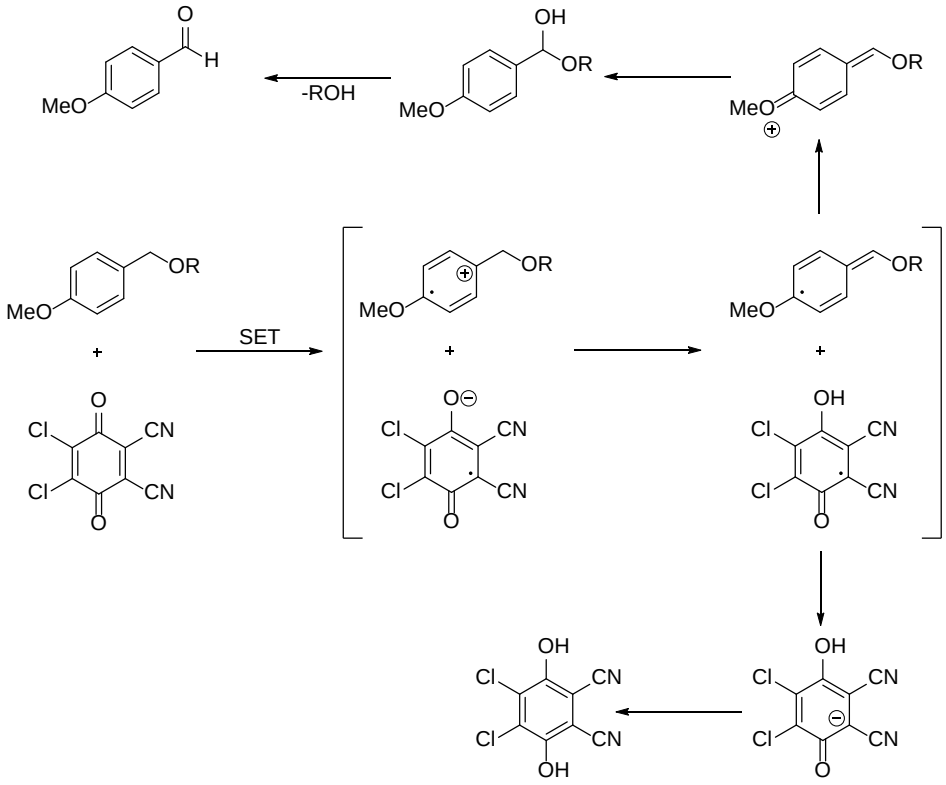
Scheme 2-7: Diol **90** was prepared by temporary acetal protection of **88** followed by PMB protection and acetal deprotection.

Quantitative benzoylation of **90** was achieved with benzoyl chloride (BzCl) in the presence of triethylamine (Et_3N) and DMAP. After ended reaction excess BzCl was removed by addition of 3-(dimethylamino)-1-propylamine (DMAPA), rendering flash column chromatography of the product unnecessary (Scheme 2-8).[116] Subsequent reaction with DDQ in a 9:1 mixture of water and CH_2Cl_2[117] provided the bifunctional product **81**, needed for linear elongation of the xylose backbone.



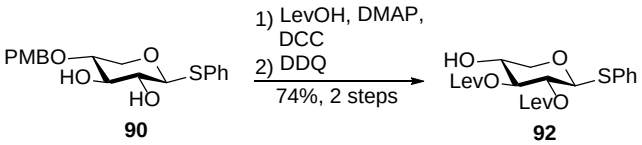
Scheme 2-8: Synthesis of phenyl 2,3-di-*O*-benzoyl-1-thio-β-D-xylopyranoside (**81**).

The DDQ deprotection is thought to happen through a single electron transfer forming an oxonium ion, which is then being neutralised by water (Scheme 2-9). DDQ is simultaneously being reduced to 2,3-dichloro-5,6-dicyanohydroquinone, which precipitates as it is not soluble in either water or CH_2Cl_2. In this way the reaction medium is kept neutral through the transformation protecting e.g. acid sensitive protection groups.[112,118]



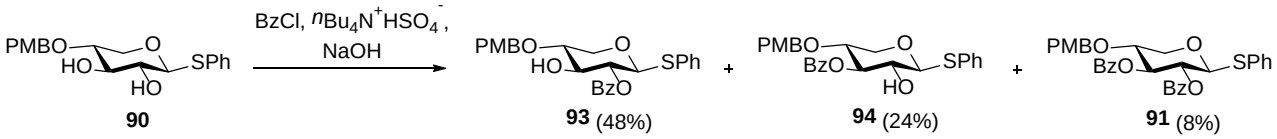
Scheme 2-9: Mechanism for DDQ deprotection of PMB.[118]

The diol **90** is furthermore the intermediate for the bifunctional product **92** with two Lev groups needed for the formation of a disubstituted xylose unit. Installation of the two Lev groups proceeded smoothly by reaction with 4-oxopentanoic acid, DMAP and *N,N′*-dicyclohexylcarbodiimide (DCC).[115] The following oxidative cleavage of the PMB group by DDQ provided **92** (Scheme 2-10).



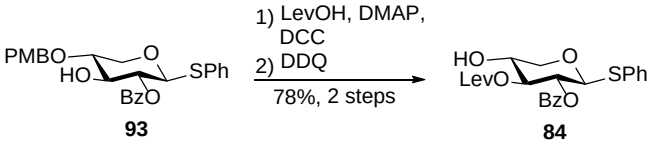
Scheme 2-10: Formation of phenyl 2,3-di-*O*-levulinoyl-1-thio-β-D-xylopyranoside (**92**).

For the monomeric xylose building block representing the *O*-3 Ara*f* monosubstituted xylose unit in the backbone, benzoyl protection of the *O*-2 position and Lev protection of the *O*-3 position are needed. The regioselective benzoylation at *C*-2 was achieved by employing the phase-transfer agent tetrabutylammonium hydrogen sulfate[119] (Scheme 2-11). The regioselectivity goes against the general established reactivity of the hydroxyls of β-xylosides, where an explanation could be found in the well-documented tendency of Bz groups to migrate especially in base.[120,121] Hence, the ester is initially formed at the more reactive *O*-3 position, but then migrates to *O*-2 position. The desired product **93** can be isolated as the major product (48%) together with both the *O*-3 benzoylated (**94**) (24%) and the di-benzoylated (**91**) products (8%), where both by-products can be used for either synthesis of the building block for monosubstituted *O*-2 Ara*f* xylose unit (*vide infra*) or **81**, respectively.



Scheme 2-11: Regioselective benzoylation of **90** through the use of a phase transfer agent forming **93** as the major product.

Subsequent installation of a Lev group at the *O*-3 position and DDQ deprotection following the same procedures as previously afforded the last bifunctional product **84** for the synthesis of the xylose backbone for this project (Scheme 2-12).

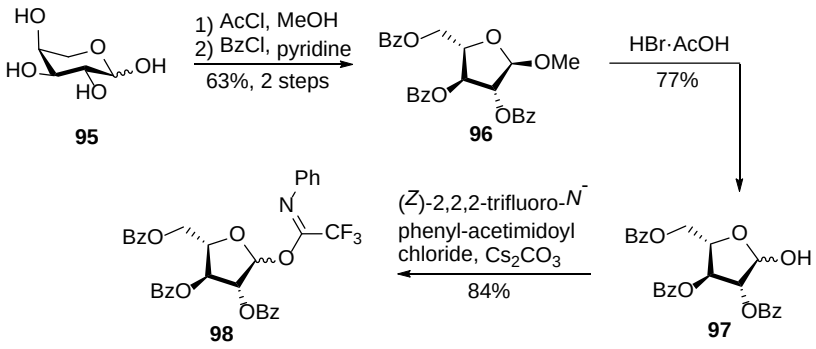


Scheme 2-12: Synthesis of phenyl 2-*O*-benzoyl-3-*O*-levulinoyl-1-thio-β-D-xylopyranoside (**84**).

2.4 Synthesis of monomeric arabinose building block

The arabinose building block needed for the construction of the targets was envisioned as a perbenzoylated substrate, with the possibility of neighbouring group participation, and as the

N-phenyl trifluoroacetimidate donor functionality due to previous experience in the department.[122] The donor was synthesised in four steps from commercially available L-arabinose (**95**). Firstly, the methyl glycoside was formed followed by benzoylation.[123] Acetolysis of **96** using hydrogen bromide solution in acetic acid[124] formed the hemiacetal **97**. During the studies, it was found the acetolysis using HBr·AcOH provided a higher yield and a shorter reaction time than when using trifluoroacetic acid. Subsequently the hemiacetal **97** was converted into the acetimidate **98** by reaction with (*Z*)-2,2,2-trifluoro-*N*-phenyl-acetimidoyl chloride and cesium carbonate.[122,125] The utilisation of the cesium salt provided a cleaner reaction and higher yield than when the potassium salt was used. Acetimidate **98** was formed as an anomeric mixture of which complete separation was not possible, but also not necessary either since both anomers could be used for the later glycosylations. The acetimidate donor could be stored in the freezer over an extended period of time without degradation.



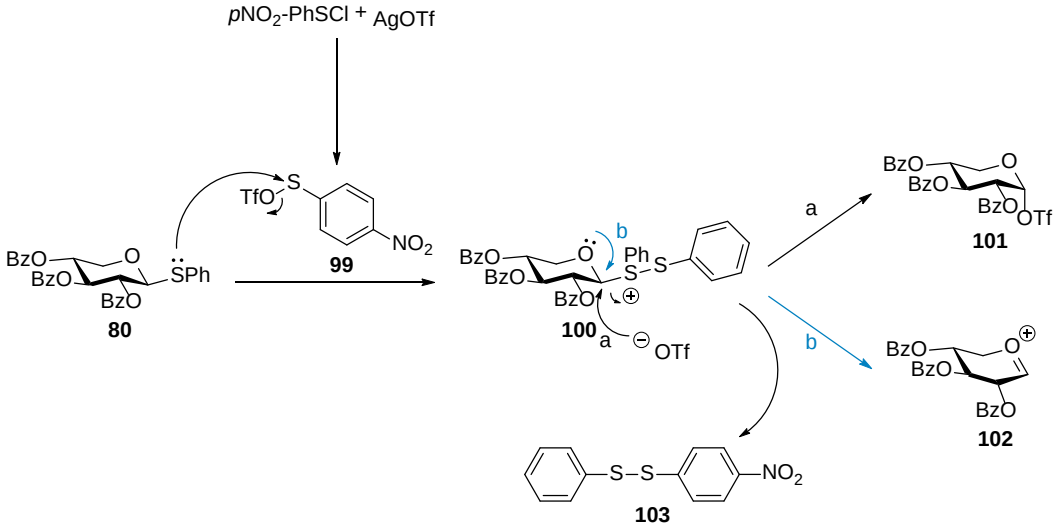
Scheme 2-13: Synthesis of the 2,3,5-tri-*O*-benzoyl-α/β-L-arabinofuranosyl 1-(*N*-phenyl)-2,2,2-trifluoroacetamidate donor **98**.

2.5 Assembly of protected AX target structures

As described above in section 2.2 a novel linear, iterative synthetic strategy for the xylose backbone assembly has been developed in our laboratories. This relies on the pre-activation of a single thioglycosyl donor functionality promoted by the commercially available *p*NO_2-PhSCl, which had been sublimated, and subsequent addition of the acceptor, where repetition of the aforementioned steps forms the linear backbone.[105]

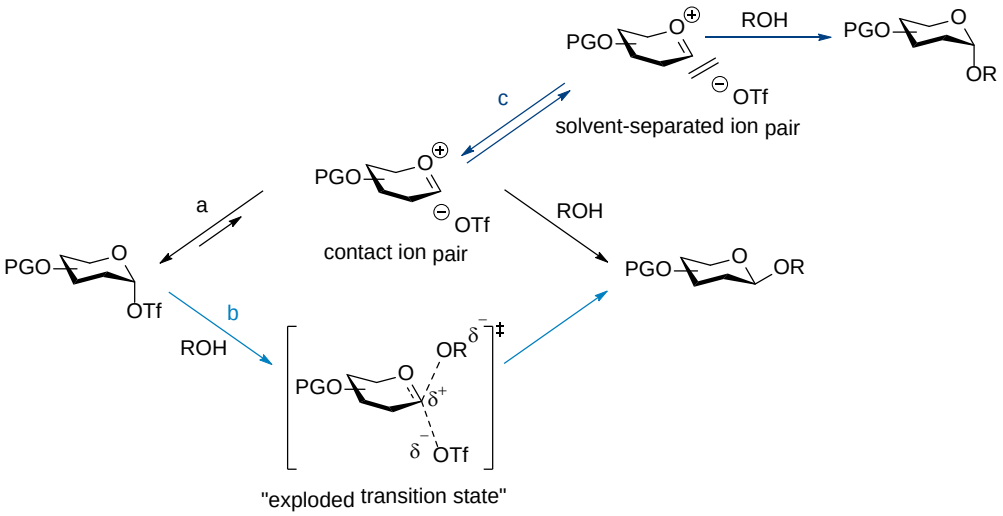
The actual activation of the thiodonors occurs by the means of *p*NO_2-PhSOTf (**99**) that is generated *in situ* at -65 °C by reaction of the corresponding chloride and the silver triflate, which is dissolved in dichloromethane and toluene, respectively. The plausible mechanism for activation of the thiodonors, here exemplified by **80**, is presented in Scheme 2-14. Two possible routes can be followed after formation of the disulfide bond, leading to either the α-triflate 101

or the oxycarbenium ion **102**. Both pathways trigger the departure of the disulphide **103** as a coloured and insoluble byproduct.



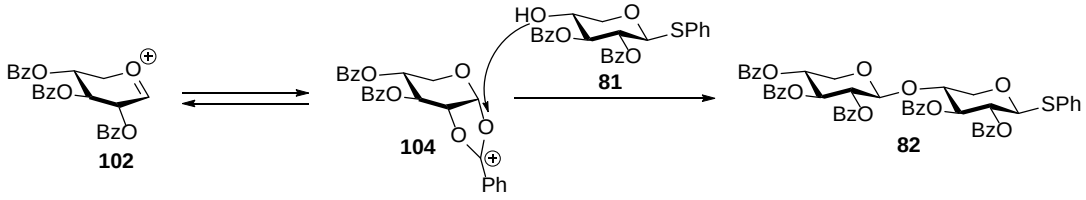
Scheme 2-14: Proposed mechanism for the activation of the donor by the *p*NO_2-PhSCl/AgOTf promotor system by formation of the α-triflate **101** (a) or the oxycarbenium ion **102**(b).

Several studies have confirmed the formation of the α-triflate intermediate at -78 °C (Scheme 2-14, pathway a).[126–128] The intermediate has never been isolated as it degrades rapidly at elevated temperatures, but it has been observed using low-temperature NMR experiments for studying triflate-mediated glycosylations.[126–128] The proposed reaction progress involves the formation of the highly electrophilic sulfenyl triflate, corresponding to **99**, which reacts so rapidly with the sulfur atom of the donor, that it cannot be observed by NMR.[126,127] Then, the covalent α-triflate is formed giving rise to the β-selectivity in the S_N2-like displacement reaction with glucose based donors.[127] However, in 2004 Crich and Chandrasekera[129] argued that as the stereoselectivity is often not complete, then the selectivity cannot be caused by a pure S_N2 mechanism as the covalent triflate indicates. Through measuring the kinetic isotope effect, they concluded that the displacement of the α-triflate occurs with substantial oxycarbenium character. This was explained by the formation of a transient contact ion pair where the triflate ion shields the α-face (Scheme 2-15 pathway a) or through an "exploded transition state" with a loose association of the nucleophile to the anomeric center while the leaving groups departs (Scheme 2-15 pathway b). The equilibrium between the contact ion pair and the solvent separated ion pair could give rise to the minor α-glycoside often obtained (Scheme 2-15 pathway c). However, the exact mechanism is still being discussed.[128,130]



Scheme 2-15: Simplified mechanism hypothesis of stereoselectivity using triflates.[129] Formation of a contact ion pair, shielding the α-face (a) or a loosely associated "exploded transition state (b) both leading to β-glycosylation. α-glycosylation can occur through the equilibrium of a contact ion pair and a solvent-separated ion pair (c).[129]

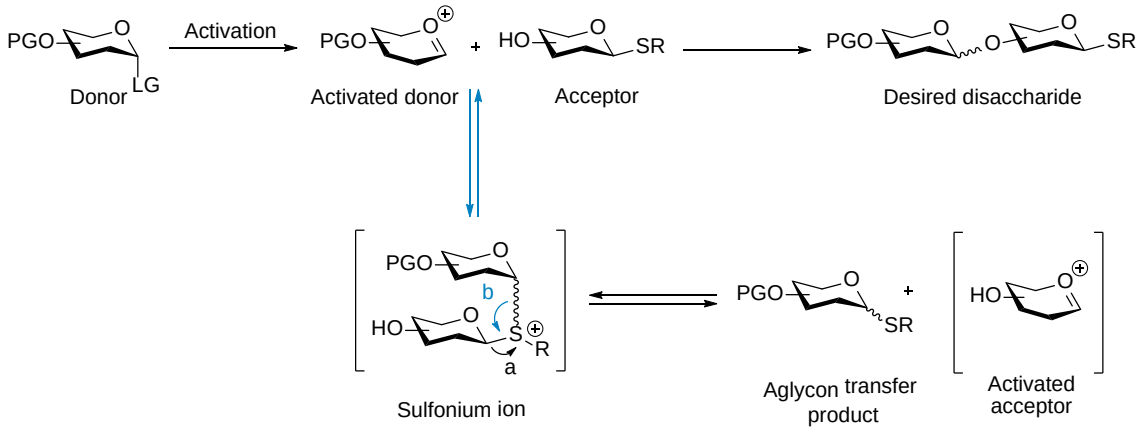
The alternative mechanism from Scheme 2-14 (pathway b) is the direct formation of the oxycarbenium ion **102**, which would be stabilised by the presence of the participating *C*-2 benzoyl group generating the acyloxonium ion **104** also giving rise to the β-product, here exemplified by reaction with **81** (Scheme 2-16).[131] Hence several explanations can be given for the observed β-selectivity of the glycosylations.



Scheme 2-16: Glycosylation of the acyloxonium ion.

When employing thioglycosides as acceptors the most important side reaction is aglycone transfer, which is generally caused by competition between the sulfur and the hydroxyl group on the acceptor for the activated oxycarbenium ion of the donor.[132–134] The intermolecular aglycone transfer is largely considered to happen through the formation of a di-glycosyl sulfonium ion by reaction between the sulfur atom on the thioglycoside acceptor and the activated glycosyl donor (Scheme 2-17). Subsequent break down of the sulfonium ion can happen in two ways. Either transfer of the aglycone to the donor by cleavage of the bond between the sulfur atom and the anomeric carbon on the acceptor (pathway a). Otherwise

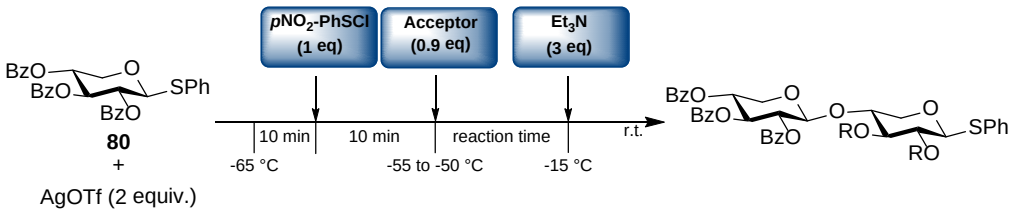
cleavage of the bond between the anomeric carbon on the donor and the sulfur will reform the original acceptor (pathway b), making formation of the desired disaccharide possible.[133,134]



Scheme 2-17: Proposed mechanism for aglycon transfer of thioglycosides. Formation of the sulfonium ion can lead either to aglycone transfer and activation of the acceptor (a) or back to the activated donor (b) and the desired disaccharide.[133,134]

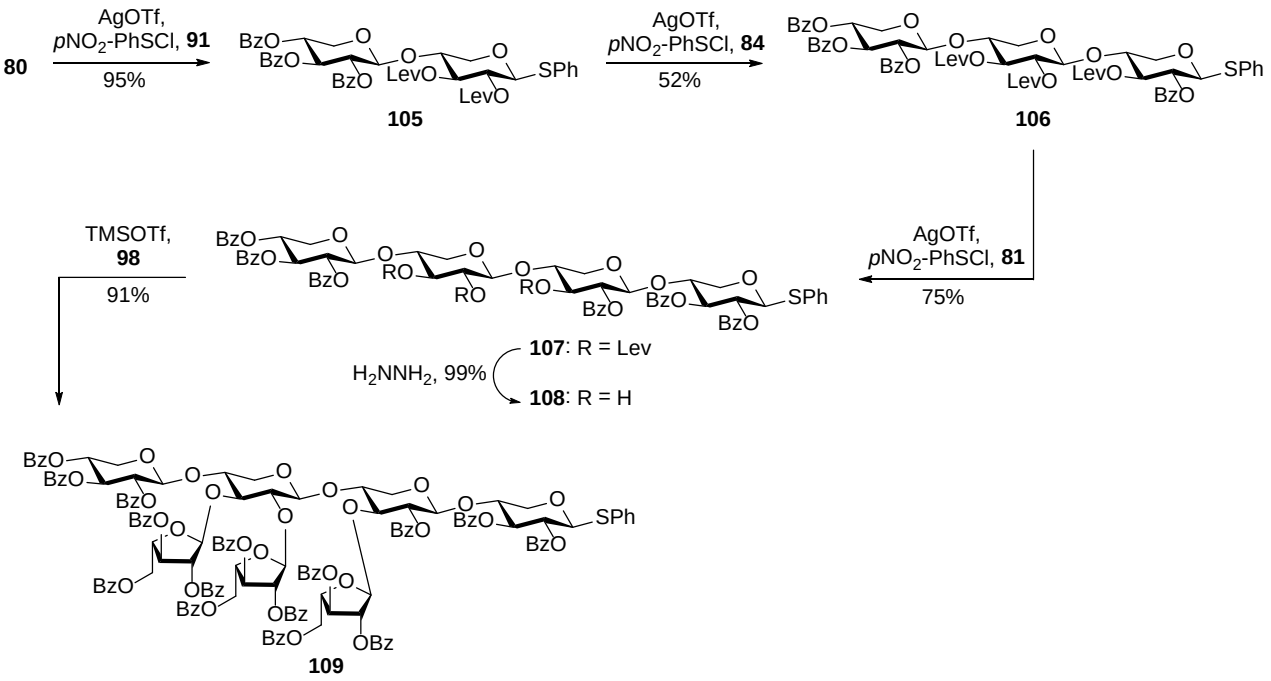
Aglycon transfer mainly happens if the hydroxyl group on at acceptor has a lower reactivity compared to the sulfur, although many factors can influence the prevalence of the aglycone transfer e.g. temperature, sterics, electronics, solvents, and donor functionality. One example that can reduce aglycone transfer through sterics is the use of the bulky 2,6-dimethylthiophenyl functionality, which still can act as a donor.[132] Relevant for this project M. F. Böhm observed that by raising the temperature for glycosylation to more than -60 °C reduced the aglycone transfer.[105] Overall the general conclusion being that the prediction of aglycone transfer is difficult even for experienced carbohydrate chemists.

The illustration of Scheme 2-18 shows the adopted glycosylation protocol with **80** as the donor. Premixing of the donor and silver triflate at -65 °C is followed by activation of the donor by addition of a stoichiometric amount of the sublimated promotor. The activation was followed by thin layer chromatography (TLC), and generally, the activation time was approximately 10 min, although for some donors additional time was needed (*vide infra*). After activation, the 4-hydroxyl acceptor is added and the temperature is raised to between -55 and -50 °C until TLC shows full conversion of the acceptor. After completion the reaction mixture is allowed to heat to -15 °C, at which point it is quenched by addition of triethylamine, exclusively providing the β-(1→4)-linked products.



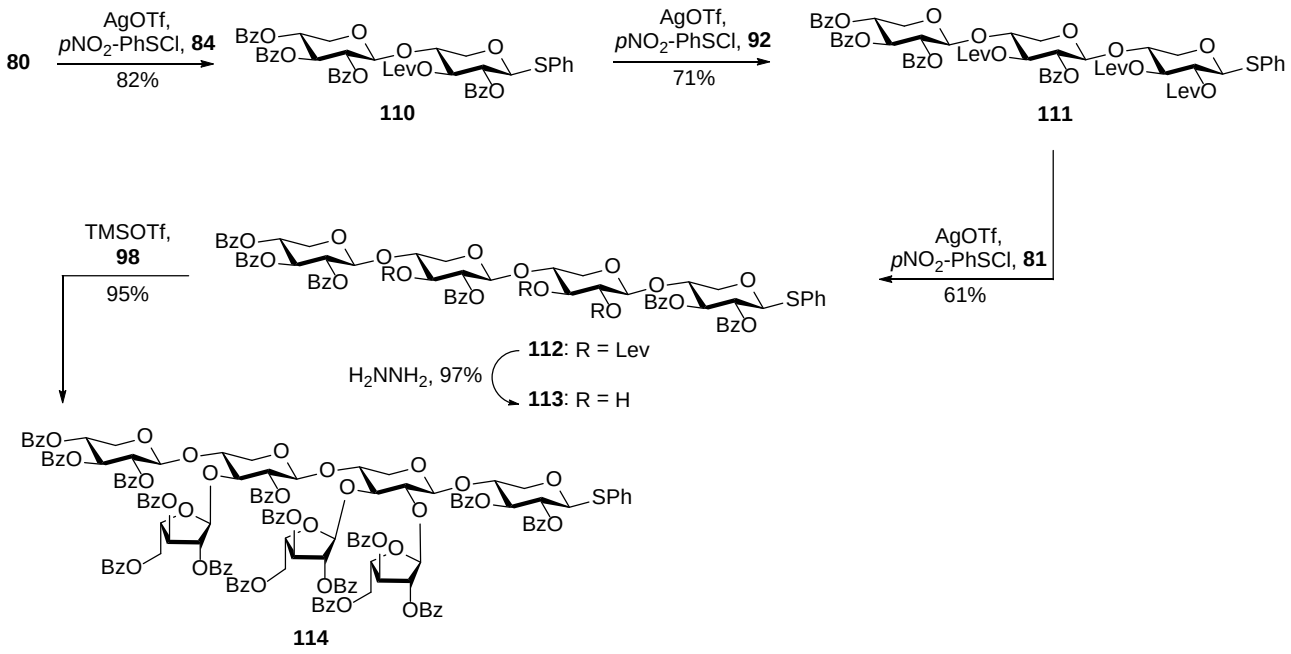
Scheme 2-18: Generalised time-scaled representation of the glycosylation protocol starting from **80**. eq: equivalents.

The first step for the synthesis of the heptasaccharide corresponding to target structure **72** was the coupling of **80** and **92**, forming disaccharide **105** in 95% yield following the aforementioned glycosylation protocol (Scheme 2-19). Next, the trisaccharide **106** was formed in 52% yield, where the donor was reisolated in 10% yield due to aglycone transfer. It is speculated, that the Lev group in the 2-position is not as prone to act as a participating neighbouring group as the benzoyl, and thus leading to a lower yield. The tetrasaccharide **106** was formed in 75% yield, followed by quantitative deprotection of the Lev groups with hydrazine hydrate under standard conditions.[135] The deprotection provided the product **108** in high yield but the reaction also had the benefit of easy work-up thereby obtaining the product without the need for chromatography. The three arabinose units were installed at -40 °C through reaction of the trifluoroactimidate donor **98** and acceptor **108** activated by TMSOTf. The reaction went smoothly with complete α-selectivity when the promotor was added as a freshly prepared stock solution in dichloromethane affording the desired heptasaccharide **109**. Addition of TMSOTf via a Hammett syringe mainly provided the heptasaccharide **109,** but also a significant amount of the hexasaccharide with the arabinoses α-(1→2)- and α-(1→3)-linked at the same xylose unit was formed (confirmed by NMR, see experimental). A probable cause might be addition of an insufficient amount of the promotor even though only a catalytic amount (0.1 equiv.) is needed.



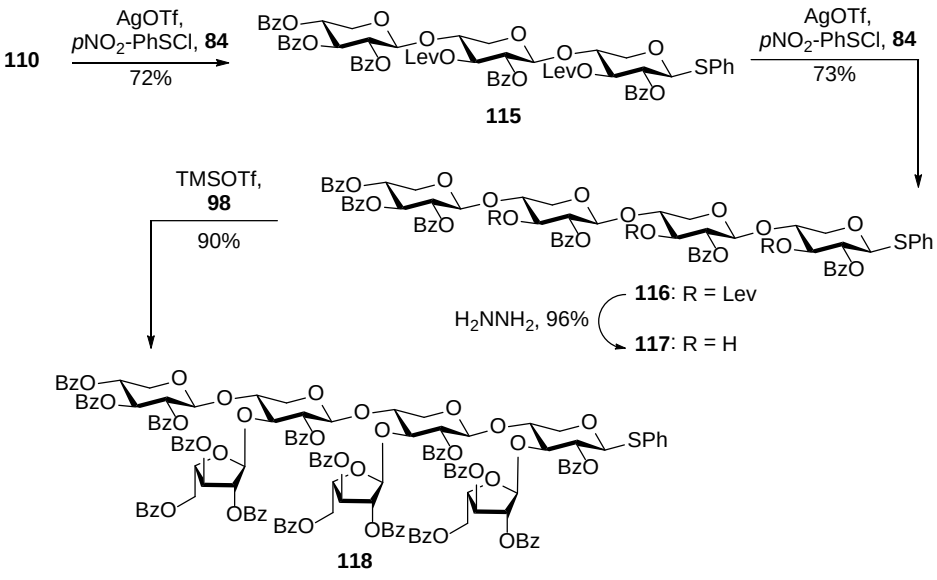
Scheme 2-19: Synthesis of heptasaccharide **109**.

The heptasaccharide corresponding to target structure **73** was assembled in a similar way (Scheme 2-20). When preforming the reactions it was observed, that disaccharide **110** needed longer time to be activated, i.e. 40 min compared to approximately 10 min of the other donors. Furthermore, the comparable lower yield for formation of the tetrasaccharide **112** (61%) mirrored the observed influence of the Lev group at the *O*-2 position from the first target structure.



Scheme 2-20: Sequential reactions affording heptasaccharide **114**.

Heptasaccharide **118** was formed utilising the same glycosylation protocol as previously starting from disaccharide **110** (Scheme 2-21). Generally, good yields were achieved for all the glycosylations, probably affected by that none of the donors had a Lev group at the *O*-2 position.



Scheme 2-21: Stepwise synthesis of heptasaccharide **118**.

During several of the glycosylations for construction of the xylose backbones approximately 10% of the donors were also isolated, indicating that aglycone transfer was not avoided completely even with strict temperature control. This also further complicated the

purifications, as the donor and the product of the respective glycosylation often presented similar R_f values. Generally, it was found that a three component solvent mixture for elution with hexane/heptane, EtOAc/acetone, and toluene gave the best separation. Especially the addition of toluene had a great influence on the separation according to TLC.

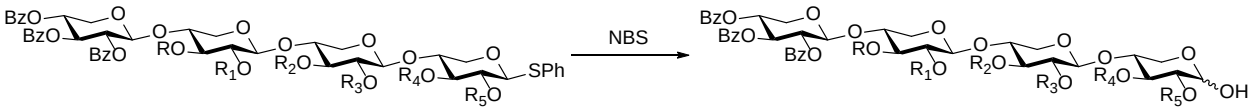
The oligosaccharide structures were analysed by NMR, and except for some signals full assignment of all 1D ^{1}H and ^{13}C spectra were achieved by the use of 2D spectra, i.e. double quantum filtered correlation spectroscopy (DFQ-COSY), heteronuclear single-quantum correlation (HSQC), heteronuclear multiple-bond correlation (HMBC), heteronuclear two-bond correlation (H2BC), and total correlation spectroscopy (TOCSY).

For several of the oligosaccharides small anomeric J_{H-H} coupling constants of around 5 Hz were observed for the (1→4)-linked xylose units. The presence of lower couplings constants usually indicate α-linkages, but the β-linkages of the xylose backbones were confirmed for either the protected or the partially deprotected tetrasaccharides by measurements of the J_{C-H} couplings constants. For **107** the couplings constants were measured to be between 165 and 159 Hz, for **113** and **117** measurements between 163-159 Hz and 163-155 Hz were observed, respectively. Coupling constants around 160 Hz is characteristic for the β-coupling (lower value are expected for the reducing end thioglycoside due to the electronegativity of the aglycon) while higher values of around 170 Hz would indicate an α-relationship.[136]

The α-relationship of the L-arabinofuranose units was concluded on the basis of the ^{1}H and ^{13}C chemical shifts as well as the very small J_{H-H} couplings for the anomeric protons.[122,137] The anomeric carbons of the α-Ara*f* units in the synthesised structures resonated around 105-106 ppm, and the anomeric protons resonated in the range of 5.25-5.76 ppm. Moreover it was observed, that the anomeric proton shift of single substituted α-(1→3)-Ara*f* unit was shifted upfield compared to when the Ara*f* units were attached to the same xylose unit. Furthermore, the anomeric protons of the Ara*f* units presented themselves as singlets indicating very small J_{H-H} couplings characteristic of α-anomers (α-anomer: 0-2 Hz, β-anomer: 4-5 Hz). Based on these observations the correct stereoselectivity was assumed, this was further corroborated by the fact that only one anomer is obtained in the reactions, and the participating effect of the benzoyl group at *O*-2 would strongly favour the α-anomer.

2.6 Deprotection of AX target structures

Following the synthesis of the three fully protected heptasaccharides **109**, **114**, and **118**, the structures were subjected to deprotection. The first strategy for deprotection involved the deprotection of the thiophenyl moiety at the reducing end followed by removal of the benzoyl groups. The hydrolysis of the anomeric thiophenyl moiety by *N*-bromosuccinimide (NBS) in a mixture of acetone and water proceeded smoothly for all three structures in good yields (Scheme 2-22). The products were isolated as α/β-mixtures, with the major product being the α-anomers, whose structures were corroborated by NMR assignments. The fact that NMR shows the α-anomers as the major products could be due to the anomer effect.

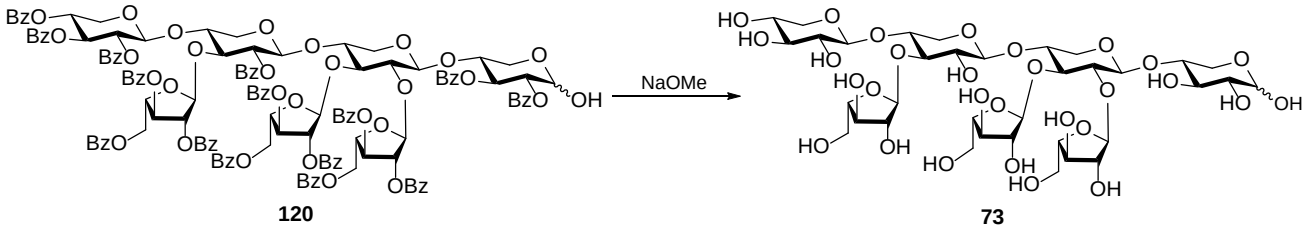


109: R, R_2 = perbenzoylated $^{\alpha}$-(1,3)-L-Ara*f*, R_1 = perbenzoylated $^{\alpha}$-(1,2)-L-Ara*f*, R_3, R_4, R_5 = Bz
114: R, R_2 = perbenzoylated $^{\alpha}$-(1,3)-L-Ara*f*, R_3 = perbenzoylated $^{\alpha}$-(1,2)-L-Ara*f*, R_1, R_4, R_5 = Bz
118: R, R_2, R_4 = perbenzoylated $^{\alpha}$-(1,3)-L-Ara*f*, R_1, R_3, R_5 = Bz

119: R, R_2 = perbenzoylated $^{\alpha}$-(1,3)-L-Ara*f*, R_1 = perbenzoylated $^{\alpha}$-(1,2)-L-Ara*f*, R_3, R_4, R_5 = Bz, (92%)
120: R, R_2 = perbenzoylated $^{\alpha}$-(1,3)-L-Ara*f*, R_3 = perbenzoylated $^{\alpha}$-(1,2)-L-Ara*f*, R_1, R_4, R_5 = Bz, (93%)
121: R, R_2, R_4 = perbenzoylated $^{\alpha}$-(1,3)-L-Ara*f*, R_1, R_3, R_5 = Bz, (92%)

Scheme 2-22: Deprotection of the anomeric thiophenyl group using NBS giving α/β mixtures of the products.

Following the deprotection of the thiophenyl moiety, the cleavage of the benzoyl groups was attempted mainly using **120** as the substrate for the transesterification. Based on previous observations from our labs the benzoyl group removal was attempted with the Zemplén conditions using a freshly prepared 1 M NaOMe solution (approximately 20 equiv.) which was added to a solution of the starting material in MeOH and dichloromethane 1:1 and left to react for 4-7 days (Scheme 2-23).



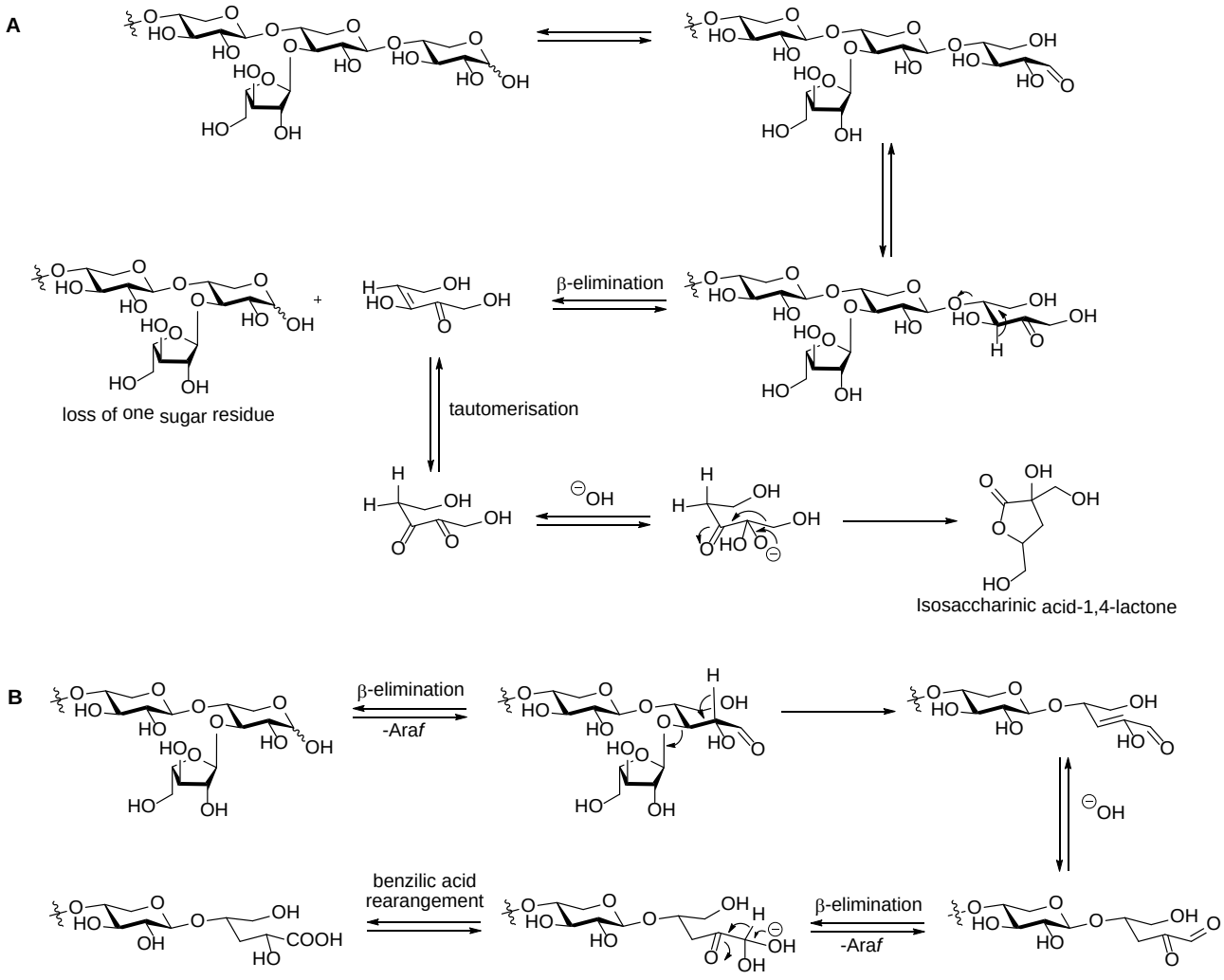
120 **73**

Scheme 2-23: Removal of the benzoyl groups under Zemplén conditions.

The reaction was challenging to follow by TLC due to the high polarity of the product. However, when NMR spectra were obtained after neutralisation by addition of Amberlite IR-120 H^+ form, it was seen that a mixture of compounds were formed. Additionally, it was found that a reaction time of 7 days was needed for complete the deprotection of the benzoyl groups. As a mixture of

compounds were obtained, it was attempted to purify these using several different methods. Separation by flash column chromatography, PD MiniTrap G-10 (from GE Healthcare, separation based on difference in size), and semi-preparative HPLC with evaporative light scattering (ELS) detection were tried without much success. Manual reverse phase C18 columns (Sep Pak C18 from Waters, separation based on hydrophobicity) was tried, and by using this method it was found that it was possible to separate partially deprotected structures and fully deprotected structures. It was attempted to remove possible apolar byproducts by precipitation of the target structure in a mixture of MeOH and diethyl ether, this technique was tried both alone and in combination with one of the chromatographic methods, but no pure products were obtained.

After several trials it was found that the reaction conditions for the benzoyl cleavage was probably too harsh, leading to decomposition of the product by so-called peeling. The peeling reaction occurs for glucomannan, xylose and cellulose structures. The hemiacetal form is in equilibrium with the open-ring aldehyde form, which in alkaline medium furthermore is in equilibrium with the keto form, from which β-elimination can lead to the loss of one sugar unit (Scheme 2-24A). AX structures however is less prone to peeling if a 3-linked Ara*f* unit is present due to the so-called stop-reaction, which can lead to a more stable compound (Scheme 2-24B).[138,139] The conclusion that peeling was happening in this case was based on the previously mentioned NMR analysis and experience obtained by the Food Chemistry group at Wageningen University and Research. In their laboratories, it was found that arabinoxylan structures isolated from natural sources undergo peeling during alkaline pretreatment, where the medium is much less basic than the Zemplén conditions tried for this project (4% NaOH compared to 1M NaOMe).[138–140]



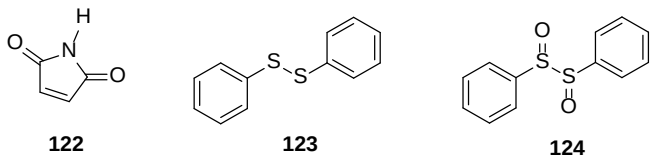
Scheme 2-24: A: Proposed peeling mechanism during alkaline treatment leading to loss of one sugar unit. B: Stop-reaction leading to stabilisation due to the presence of a 3-linked Ara*f* unit.[138,139]

During previous trials, further observations revealed that even though mixtures of unprotected sugars were obtained, the mass balance was not good. It was speculated that the target structures might adhere to the Amberlite IR-120 H^+ form during neutralisation and were not being released during washing of the beads with water. The same happened if Amberlite MB20 mixed bed resin was used. Consequently, it was attempted to neutralise the reaction mixture by addition of ammonium chloride, but unfortunately, this gave rise to large amounts of salts. Hence, it was decided to neutralise the reaction mixtures by addition of 0.1 M HCl for future trials.

Based on the previous experiences, it was decided to reverse the order of the deprotections. When the fully protected heptasaccharide **118** was subjected to Zemplén conditions (dry CH_2Cl_2 and MeOH 1:1, NaOMe added until basic, and the reaction neutralised by 0.1 M HCl after

complete reaction), it was found that the reaction went smoothly and full conversion was achieved after only 4 h compared to 7 days before. It was observed that the product was very sensitive towards acidic conditions, and breakdown of the structure was observed if pH less than 5 was achieved during neutralisation, which can be explained by the acid sensitivity of the α-linkage to the Ara*f* units. After neutralisation, the reaction mixture was just washed with Et_2O and EtOAc and dried. From NMR and MS it was seen that the expected product had been obtained, although probably some residual salts were present based on the yield (151%). Hence, further purification is needed for future reactions.

Due to time restraints, the removal of the anomeric thiophenyl moiety was only tried on a mixture of debenzoylated heptasaccharide **114** and an unknown hexasaccharide, which was formed due to cleavage of one arabinose unit under too acidic conditions during neutralisation. However, this mixture was used as trial substrate in order to preserve valuable material. The reaction was run in a mixture of acetone and water 3:1 due to solubility issues of the mixture in acetone. NBS (8.5 equiv.) was added to the reaction, which could be followed by TLC. After ended reaction (1.3 h), the mixture was concentrated, dissolved in H_2O and washed with EtOAc to remove organic residues. Afterwards it was first tried to purify the crude product by PD MiniTrap G-10. From ^{1}H NMR it could be seen, that the thiophenyl moiety had been removed, but two other molecules eluted together with the deprotected sugars. Hence the collected fractions were instead subjected to purification by reverse phase C18 columns (Sep Pak C18 from Waters), but the two previously impurities co-eluted with the sugars again. One of these molecules were identified as succimide (**122**), and the other as either diphenyl disulphide (**123**) or diphenyl disulfoxide (**124**) based on ^{1}H NMR (Figure 2-4).[141,142] However for **123** and **124** the exact identity is difficult to ascertain based on ^{1}H NMR, but for a later experiment ^{13}C NMR and MS analysis would be a possibility.



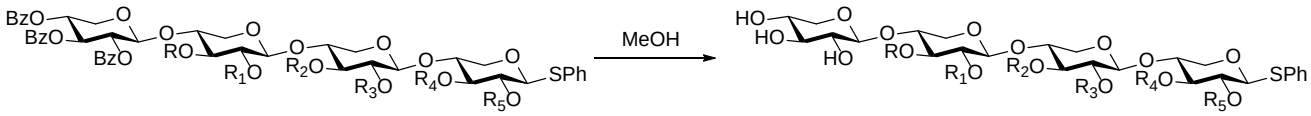
122 **123** **124**

Figure 2-4: Possible structures of co-eluting molecules.

2.7 Future work

The time limitations of this project meant that the deprotections of the obtained heptasaccharides **109**, **114**, and **118** were not accomplished at the deadline of this thesis. However, it is the hope that the syntheses will be finalised shortly.

The removal of the benzoyl groups of the other heptasaccharides should be possible using the same conditions as for compound **118** (Scheme 2-25). The first try for purification of the deprotected compound will be using a PD MiniTrap G-10 column, as this is a desalting method and the problem seems to be the presence of salts. It would, however, be preferable to subject the crude product from the debenzoylation directly to the following deprotection of the thiophenyl moiety and then only perform one purification after this final step.



109: R, R_2 = perbenzoylated $^{\alpha}$-(1,3)-L-Ara*f*, R_1 = perbenzoylated $^{\alpha}$-(1,2)-L-Ara*f*, R_3, R_4, R_5 = Bz
114: R, R_2 = perbenzoylated $^{\alpha}$-(1,3)-L-Ara*f*, R_3 = perbenzoylated $^{\alpha}$-(1,2)-L-Ara*f*, R_1, R_4, R_5 = Bz
118: R, R_2, R_4 = perbenzoylated $^{\alpha}$-(1,3)-L-Ara*f*, R_1, R_3, R_5 = Bz

125: R, R_2 = $^{\alpha}$-(1,3)-L-Ara*f*, R_1 = $^{\alpha}$-(1,2)-L-Ara*f*, R_3, R_4, R_5 = H
126: R, R_2 = $^{\alpha}$-(1,3)-L-Ara*f*, R_3 = $^{\alpha}$-(1,2)-L-Ara*f*, R_1, R_4, R_5 = H
127: R, R_2, R_4 = $^{\alpha}$-(1,3)-L-Ara*f*, R_1, R_3, R_5 = H

Scheme 2-25: Future deprotection of the heptasaccharides.

Several different solutions to avoid the observed problems with deprotection of the thiophenyl moiety have been envisioned. Firstly it would be interesting, if the deprotection could be successful using less equivalents of NBS.[143] This would limit the amount of organic by-products as well as the risk of obtaining HBr in the reaction mixture, causing decomposition of the target. Furthermore, it is speculated if the amounts of the co-eluting by-products can be decreased by precipitation of the target structure in a mixture of MeOH and diethyl ether prior to other types of purification.

2.8 Concluding remarks

For this project, three arabinoxylan fragments were envisioned as the targets. To accomplish the synthesis four different xylopyranose building blocks and one arabinofuranose building block were synthesised. The xylopyranose building blocks were formed through a divergent synthesis from D-xylopyranose (**87**), and 4-*O*-PMB protected thioxylopyranoside **90** functioned as a common intermediate, from which three of the building blocks all could be formed in two to three synthetic steps (**81**, **84**, and **92**).

The xylopyranose building blocks were coupled to form the xylose backbones through a novel linear, iterative synthetic strategy, which had previously been developed in the group. Then partial deprotection of the orthogonal Lev protection group made the fully stereoselective coupling with the arabinofuranose building block possible, hence leading to the three fully protected heptasaccharides **109**, **114**, and **118** in 33%, 33%, and 37% overall yield, respectively. Through the glycosylations it was found that the presence of a Lev group at the 2-position led to lower yields in the glycosylations, which may be caused by the Lev group being less prone to act as a participation neighbouring group.

After several trials with deprotections, it was found that the benzoyl groups of heptasaccharide **118** could be removed although further purification of the product **127** is needed. This protocol is expected to be directly transferrable to the other targets. Furthermore, a test reaction for the removal of the thiophenyl moiety has been run, where the preliminary results are promising. Consequently, the syntheses of the three arabinoxylan fragment are expected to be finished in due course.

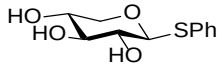
2.9 Experimental

General considerations

Starting materials, reagents, and solvents were purchased from commercial suppliers and were used without further purification unless otherwise noted. All solvents used were of analytical HPLC grade. Anhydrous solvents were obtained from an Innovative Technology PS-MD-7 PureSolv solvent purification system. Except for dichloromethane and toluene for glycosylations which were dried over 3 Å molecular sieves prior to use.[144] Air- and moisture sensitive experiments were conducted under inert atmosphere (N_2) in dried glassware and anhydrous solvents. Water content was checked by 899 Coulometer from Metrohm. Crushed 3 Å molecular sieves for glycosylations were stored in the oven at 110 °C, before activation with heating under vacuum. Evaporation of solvents was performed with a VWR International Laborota 400 under reduced pressure (*in vacuo*) at temperatures between 23-50 °C. Thin-layer chromatography (TLC) was carried out using Merck Aluminium Sheets pre-coated with 0.25 mm silica gel, C-60 F_{254} plates. TLC plates were inspected under UV light or visualized by charring after dipping in a cerium ammoniumsulfate solution (1% cerium(IV)sulphate and 2.5% ammonium heptamolybdate in a 10% sulfuric acid solution). Flash column chromatography was performed using Merck Geduran silica gel 60 Å (40-63 µm) as the stationary phase, dry column chromatography was accomplished using Merck silica gel 60 Å (0.015-0.040 mm). Eluent systems are specified for each R_f-value, and ratios are given as volume ratios. NMR spectre were recorded with a Bruker Ascend™ 400 (400 MHz and 101 MHz) with a Prodigy cyroprobe or a Bruker Ascend™ 800 (800 MHz and 200 MHz). The recorded chemical shift are reported in parts per million (δ = scale) relative to the residual solvent peak in $CDCl_3$ (δ_H = 7.26 ppm, δ_C = 77.16 ppm), methanol-*d4* (δ_H = 4.87 ppm, δ_C = 49.00 ppm), and D_2O (δ_H = 4.79 ppm). The annotation (2C), (3C) ect. indicates when multiple carbon signals are overlapping. Coupling constants (*J*) are reported in Hz. Multiplicities are reported as singlet (s), broad singlet (bs), doublet (d), doublets of doublets (dd), doublets of triplets (dt), doublets of doublets of doublets (ddd), triplet (t), triplets of doublets (td), quartet (q), and multiplet (m). Assignment of ^1H and ^{13}C resonances were based on APT, DQF-COSY, HSQC, H2BC, HMBC, TOCSY, and HSQC-TOSCY when needed. Optical rotation was measured on a Perkin Elmer Model 341 Polarimeter. HRMS analysis was performed on either a UHPLC-QTOF system (Dionex ultimate 3000 and Bruker MaXis) with an electrospray ionization (ESI) source or a MALDI-TOF system (Bruker Solarix XR 7T) and controlled using Data Analysis 4.2 software.

MS of partial deprotected heptasaccharide were done using a Ultimate 3000 high performance liquid chromatography (HPLC) system coupled to electrospray ionization (ESI) with a LTQ-Velos ion trap mass spectrometer run in negative mode (Thermo Scientific). Analytes were separated using an BEH-glycan column (150 x 2.1 mm, particle size 1.7 µm) (Waters). The binary mobile phase consisted of ammonium formate aqueous solution (50 mM, pH 4.4) and acetonitrile. Data were processed using Xcalibur (Thermo Scientific) software. Elemental analysis were performed at the Microanalytic Laboratory Kolbe in Mülheim an der Ruhr (Germany). New compounds have been characterised by R_f, $[\alpha]_D$, ^{1}H NMR, ^{13}C NMR, and HRMS.

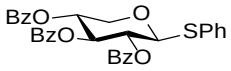
Phenyl 1-thio-β-D-xylopyranoside (88)



D-Xylose (**87**) (29.918 g, 199.28 mmol, 1.0 equiv.), Et_3N (223.0 mL, 1599.89 mmol, 8.0 equiv.) and DMAP (4.903 g, 40.13 mmol, 0.20 equiv.) were suspended in CH_2Cl_2 (300 mL) at 0 °C. Upon dropwise addition of Ac_2O (94.0 mL, 996.27 mmol, 5.0 equiv.) the solution turned yellow. After stirring for 2 h at r.t. TLC analysis (hexane/EtOAc 1:1) showed the consumption of the starting material and formation of products. The reaction mixture was washed with 1 M HCl and brine followed by concentration *in vacuo*. The crude peracetylated xylopyranoside was dissolved in dry CH_2Cl_2 (160.0 mL) followed by addition of PhSH (25.0 mL, 242.80 mmol, 1.2 equiv.) and $BF_3{\cdot}OEt_2$ (76.0 mL, 599.74 mmol, 3.0 equiv.) at 0 °C under N_2 atm. The reaction mixture was stirred overnight at r.t. where it turned purple. TLC analysis (hexane/EtOAc 2:1) showed the consumption of the starting material and formation of product. The solution was diluted with CH_2Cl_2, and washed with aq. sat. $NaHCO_3$ and brine, dried over Na_2SO_4, filtered, and conc. *in vacuo*. The residue was dissolved in methanol (160.0 mL) and solid sodium until basic. After 1 h TLC analysis (hexane/EtOAc 2:1) showed the consumption of the starting material and formation of product. The solution was neutralised with Amberlite IR-120 (H^+), filtered, and conc. *in vacuo*. The residue was purified by dry column chromatography (10% acetone in toluene, 5% gradient), followed by recrystallization in hexane and acetone, where the white crystalline product, **88**, was obtained (19.161 g, 40%). **R_f**(toluene/acetone 1:1) 0.27. **^{1}H NMR** (400 MHz, Methanol-d_4) δ 7.54 - 7.51 (m, 2H, ArH_{ortho}), 7.34 - 7.26 (m, 2H, ArH_{meta}, ArH_{para}), 4.56 (d, $J_{1,2}$ = 9.3 Hz, 1H, H1), 3.95 (dd, $J_{5,5'}$ = 11.3 Hz, $J_{5,4}$ = 5.2 Hz, 1H, H5), 3.51 - 3.45 (m, 1H, H4), 3.38 - 3.31 (m, 3H, H3, OH), 3.26 – 3.19 (m, 2H, H2, H5'). **^{13}C NMR** (101 MHz, Methanol-d_4) δ 134.9 (ArC_{ipso}), 133.1 (2C, ArC_{ortho}H), 129.9 (2C, ArC_{meta}H),

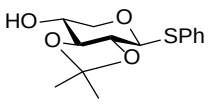
128.5 (1C, ArC_{para}H), 90.1 (C1), 79.2 (C3), 73.7 (C2), 70.9 (C4), 70.4 (C5). NMR data is in accordance with literature values.[145]

Phenyl 2,3,4-tri-*O*-benzoyl-1-thio-β-D-xylopyranoside (80)



The triol **88** (5.12 g, 21.13 mmol, 1.0 equiv.) was dissolved in pyridine (45 ml) and BzCl (7.4 ml, 63.40 mmol, 3.0 equiv.). The reaction mixture was stirred at r.t. for 1 h and the excess of BzCl was quenched by addition of methanol (10 mL) and the mixture was stirred for an additional 10 minutes. The disappearance of a white precipitate was observed. The reaction mixture was diluted with CH_2Cl_2 and washed with 1 M HCl, and water, dried over Na_2SO_4, filtered, and conc. *in vacuo*. The residue was purified by column chromatography to afford **80** (5.1 g, 84%). **R_f**(heptane/EtOAc 7:3) 0.28. **^{1}H NMR** (400 MHz, $CDCl_3$) δ 8.05 (dd, *J* = 8.2, 1.1 Hz, 2H, Ar*H*), 8.01 – 7.98 (m, 4H, Ar*H*), 7.56 – 7.51 (m, 5H, Ar*H*), 7.42 – 7.32 (m, 9H, Ar*H*), 5.78 (t, *J* = 6.6 Hz, 1H, H3), 5.46 (t, *J* = 6.3 Hz, 1H, H2), 5.33 – 5.26 (m, 2H, H1, H4), 4.71 (dd, $J_{5eq,5ax}$ = 12.3 Hz, $J_{5eq,4}$ = 4.0 Hz, 1H, $H5_{eq}$), 3.83 (dd, $J_{5ax,5eq}$ = 12.3 Hz, $J_{5ax,4}$ = 6.5 Hz, 1H, $H5_{ax}$). **^{13}C NMR** (101 MHz, $CDCl_3$) δ 165.6 (1C, C_q), 165.3 (1C, C_q), 165.3 (1C, C_q) (*C*O, Bz), 133.6 (1C, Ar*C*H), 133.5 (1C, Ar*C*H), 133.5 (1C, Ar*C*H), 133.2 (1C, Ar*C*), 132.8 (2C, Ar*C*H), 130.2 (2C, Ar*C*H), 130.1 (2C, Ar*C*H), 130.1 (2C, Ar*C*H), 129.3 (2C, Ar*C*), 129.2 (2C, Ar*C*H), 129.1 (1C, Ar*C*), 128.6 (2C, Ar*C*H), 128.6 (2C, Ar*C*H), 128.5 (2C, Ar*C*H), 128.3 (1C, Ar*C*H), 86.5 (C1), 70.6 (C3), 70.1 (C2), 68.8 (C4), 63.7 (C5). **Elemental Analysis:** calc. C: 69.30 H: 4.73 S: 5.78; found: C: 69.26 H: 4.70 S: 5.66. NMR data are in accordance with literature values.[146]

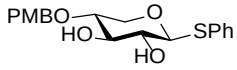
Phenyl 2,3-*O*-isopropylidene-1-thio-β-D-xylopyranoside (89)



To a solution of **88** (2.018 g, 8.32 mmol, 1.0 equiv.) in dry DMF (15.0 mL) was added CSA (0.284 g, 1.22 mmol, 0.15 equiv.) and the mixture was heated to 60 °C. 2-Methoxy prop-1-ene (2.4 mL, 24.96 mmol, 3.0 equiv.) was slowly added to the mixture. After 2.5 h of stirring additional 2-methoxy prop-1-en (1.0 mL, 18.49 mmol, 2.2 equiv.) was added. A further amount of 2-methoxy prop-1-ene (1.0 mL, 18.49 mmol, 2.2 equiv.) was added after a total reaction time of 4 h. After 4.5 h TLC analysis (toluene/AcOH 1:1) showed the consumption of the starting material and formation of two products. The reaction mixture was cooled to r.t., diluted with DCM, and washed with H_2O and brine, dried over Na_2SO_4, filtered, and conc. *in vacuo*. The residue was purified by flash column chromatography

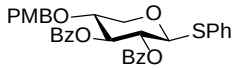
(hexane/DCM/EtOAc 4:1:1). The product was a white solid (1.187 g, 50%). **R_f**(toluene/acetone 3:1) 0.62. **^{1}H NMR** (400 MHz, $CDCl_3$) δ 7.58 – 7.55 (m, 2H, ArH_{ortho}), 7.34 – 7.30 (m, 3H, ArH_{meta}, ArH_{para}), 4.80 (d, $J_{1,2}$ = 9.6 Hz, 1H, H1), 4.13 (dd, $J_{5,5'}$ = 11.6 Hz, $J_{5,4}$ = 5.3 Hz, 1H, $H5_{eq}$), 4.03 – 3.96 (m, 1H, H4), 3.54 (t, J = 9.1 Hz, 1H, H3), 3.27 – 3.22 (m, 2H, H2, $H5_{ax}$), 2.17 (d, $J_{OH,4}$ = 4.0 Hz, 1H, OH), 1.49 (s, 3H, CCH_3), 1.45 (s, 3H, CCH_3). **^{13}C NMR** (101 MHz, $CDCl_3$) δ 133.0 (2C, ArC_{ortho}H), 132.0 (1C, ArC_{ipso}), 129.0 (2C, ArC_{meta}H), 128.3 (1C, ArC_{para}H), 111.5 (1C, $C(CH_3)_2$), 85.6 (C1), 83.0 (C3), 75.3 (C2), 70.0 (C5), 69.2 (C4), 26.8 (CCH_3), 26.7 (CCH_3). NMR data are in accordance with literature values.[147]

Phenyl 4-*O*-(*p*-methoxy)benzyl-1-thio-β-D-xylopyranoside (90)



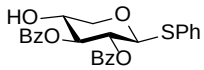
Acetonide **89** (6.414 g, 22.72 mmol, 1.0 equiv.) was dissolved in dry DMF (45.0 mL) and NaH 60% in mineral oil (1.632 g, 40.81 mmol, 1.8 equiv.) was added at 0 °C. After 10 min 4-methoxybenzyl chloride (4.0 mL, 29.37 mmol, 1.3 equiv.) was added, followed by stirring of the reaction mixture at r.t. overnight. TLC analysis (toluene/AcOH 4:1) showed the consumption of the starting material and formation of product. The reaction was quenched by addition of 2 M HCl and diluted with CH_2Cl_2. The solution was washed with aq. sat. $NaHCO_3$, dried over Na_2SO_4, filtered, and conc. *in vacuo*. After filtration the solution was diluted with MeOH (35.0 mL), CSA (5.268 g, 22.68 mmol, 1.0 equiv.) was added. The mixture was allowed to stir at r.t. After 1 h TLC analysis (hexane/EtOAc 5:1) showed the consumption of the starting material as well as the formation of product (hexane/EtOAc 1:1). The reaction was quenched by addition of Et_3N and conc. *in vacuo*. The residue was purified by flash column chromatography (hexane/EtOAc 1:1), and the product obtained as a white crystalline solid (7.217 g, 88% over 2 steps). **R_f**(hexane/EtOAc 1:1) 0.47. $[\alpha]_D^{298\,K}$ = -55.9 (*c* 1, $CHCl_3$). **^{1}H NMR** (400 MHz, $CDCl_3$) δ 7.53 – 7.50 (m, 2H, Ar$H_{ortho(SPh)}$), 7.33 – 7.29 (m, 3H, Ar$H_{meta(SPh)}$, Ar$H_{para(SPh)}$), 7.27 – 7.23 (m, 2H, Ar$H_{ortho(PMB)}$), 6.90 – 6.86 (m, 2H, Ar$H_{meta(PMB)}$), 4.61 (d, J_{gem} = 11.4 Hz, 1H, OCH_2Ph), 4.58 – 4.54 (m, 2H, H1 ($J_{1,2}$ = 8.9 Hz), OCH_2Ph), 4.07 (dd, $J_{5eq,5ax}$ = 11.5 Hz, $J_{5eq,4}$ = 4.8 Hz, 1H, $H5_{eq}$), 3.80 (s, 3H, OCH_3), 3.66 (t, J = 8.4 Hz, 1H, H3), 3.49 – 3.43 (m, 1H, H4), 3.40 (t, J = 8.6 Hz, 1H, H2), 3.27 (dd, $J_{5ax,5eq}$ = 11.5 Hz, $J_{5ax,4}$ = 9.7 Hz, 1H, $H5_{ax}$), 2.52 (bs, 2H, OH). **^{13}C NMR** (101 MHz, $CDCl_3$) δ 159.7 (1C, Ar$C_{para(PMB)}$), 132.8 (2C, ArC_{ortho}H $_{(SPh)}$), 132.3 (1C Ar$C_{ipso(SPh)}$), 130.1 (1C, Ar$C_{ipso(PMB)}$), 129.7 (2C, ArC_{ortho}H$_{(PMB)}$), 129.2 (2C, ArC_{meta}H$_{(SPh)}$), 128.3 (1C, ArC_{para}H$_{(SPh)}$), 114.2 (2C, ArC_{meta}H$_{(PMB)}$), 88.8 (C1), 76.6 (C4), 76.5 (C3), 72.8 (OCH_2Ph), 72.1 (C2), 67.1 (C5), 55.4 (OCH_3). **HRMS:** calcd. for $C_{19}H_{22}O_5SNa^+$ *m/z* 385.1080; found *m/z* 385.1090.

Phenyl 2,3-di-*O*-benzoyl-4-*O*-(*p*-methoxy)benzyl-1-thio-β-D-xylopyranoside (91)



Diol **90** (1.970 g, 5.43 mmol, 1.0 equiv.) was dissolved in CH_2Cl_2 (32,0 mL). Et_3N (2.4 mL, 17.22 mmol, 3.2 equiv.), DMAP (0.069 g, 0.56 mmol, 0.1 equiv.) and BzCl (1.9 mL, 16.36 mmol, 3.0 equiv.) were added to the solution at 0°C. After 5 min the reaction mixture was left to stir at r.t. overnight. TLC analysis (hexane/EtOAc 4:1) showed the consumption of the starting material and formation of product. 3-(Di-methylamino)-1-propylamine (2.2 mL, 17.48 mmol, 3.2 equiv.) was added and the reaction mixture was stirred for 1.5 h, at which point it was washed with 1 M HCl and brine, dried over $MgSO_4$, filtered, and conc. *in vacuo*. The white crystalline product (3.115 g, 100%) was used without further purification. $\mathbf{R_f}$(hexane/EtOAc 4:1) 0.28. $[\boldsymbol{\alpha}]_{\boldsymbol{D}}^{\boldsymbol{298\,K}}$ = +62.0 (*c* 1.0, $CHCl_3$). **^{1}H NMR** (400 MHz, $CDCl_3$) δ 8.00 – 7.95 (m, 3H, Ar*H*), 7.56 - 7.27 (m, 12H, Ar*H*), 7.15 – 7.12 (m, 2H, Ar*H*), 6.74 – 6.71 (m, 2H, Ar*H*), 5.60 (t, J = 8.0 Hz, 1H, H3), 5.34 (t, J = 8.0 Hz, 1H, H2), 5.04 (d, $J_{1,2}$ = 8.1 Hz, 1H, H1), 4.56 (d, J_{gem} = 11.9 Hz, 1H, OCH_2PhOMe), 4.52 (d, J_{gem} = 11.8 Hz, 1H, OCH_2Ph), 4.28 (dd, $J_{5eq,5ax}$ = 11.9 Hz, $J_{5eq,4}$ = 4.6 Hz, 1H, H5$_{eq}$), 3.80 – 3.76 (m, 1H, H4), 3.75 (s, 3H, OMe), 3.57 (dd, $J_{5ax,5eq}$ = 11.9 Hz, $J_{5ax,4}$ = 8.5 Hz, 1H, H5$_{ax}$). **^{13}C NMR** (101 MHz, $CDCl_3$) δ 165.7 (1C, C_q), 165.4 (1C, C_q) (*C*O, Bz), 159.5 (1C, Ar*C*H$_{para(PMB)}$), 133.4 (2C, Ar*C*H), 133.2 (1C, Ar*C*), 132.5 (2C, Ar*C*H), 130.1 (2C, Ar*C*H), 130.0 (2C, Ar*C*H), 129.7 (1C, Ar*C*), 129.7 (2C, Ar*C*H), 129.5 (1C, Ar*C*), 129.4 (1C, Ar*C*), 129.1 (2C, Ar*C*H), 128.5 (2C, Ar*C*H), 128.5 (2C, Ar*C*H), 128.1 (1C, Ar*C*H), 113.9 (2C, Ar*C*H), 87.0 (C1), 73.9 (C4), 73.6 (C3), 72.5 (O*C*H$_2$Ph), 70.6 (C2), 66.4 (C5), 55.4 (O*C*H$_3$). **HRMS:** calcd. for $C_{33}H_{30}O_7SNa^+$ *m/z* 593.1604; found *m/z* 593.1617. NMR data are in accordance to literature values.[148]

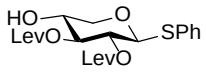
Phenyl 2,3-di-*O*-benzoyl-1-thio-β-D-xylopyranoside (81)



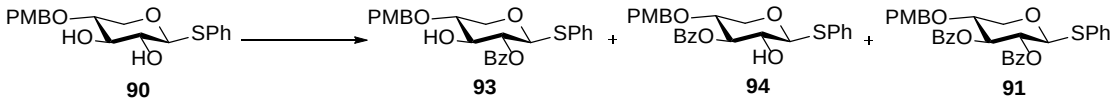
To a well stirred emulsion of **91** (0.696 g, 1.22 mmol, 1.0 equiv.) in CH_2Cl_2 (10.0 mL) and H_2O (1.0 mL), DDQ (0.415 g, 1.83 mmol, 1.5 equiv.) was added. The reaction mixture was stirred in the dark for 7 h, at which point TLC analysis (hexane/EtOAc 4:1) showed the consumption of the starting material and the formation of a product. The mixture was filtered through a big pad of Celite. The filtrate was washed with aq. sat. $NaHCO_3$ and brine, filtered and conc. *in vacuo*. Purification was done by flash column chromatography (pentane/EtOAc 4:1), and the white crystalline product (0.514 g, 94%) was obtained. $\mathbf{R_f}$(hexane/EtOAc 2:1) 0.53. $[\boldsymbol{\alpha}]_{\boldsymbol{D}}^{\boldsymbol{298\,K}}$ = +66.1 (*c* 1.0, $CHCl_3$). **^{1}H NMR** (400 MHz, $CDCl_3$) δ 8.05 –

7.99 (m, 4H, Ar*H*), 7.57 – 7.49 (m, 4H, Ar*H*), 7.43 – 7.39 (m, 4H, Ar*H*), 7.35 – 7.29 (m, 3H, Ar*H*), 5.43 (t, *J* = 7.4 Hz, 1H, H2), 5.33 (t, *J* = 7.4 Hz, 1H, H3), 5.09 (d, $J_{1,2}$ = 7.3 Hz, 1H, H1), 4.45 (dd, $J_{5eq,5ax}$ = 12.0 Hz, $J_{5eq,4}$ = 4.4 Hz, 1H, $H5_{eq}$), 4.00 (q, *J* = 6.9 Hz, 1H, H4), 3.61 (dd, $J_{5ax,5eq}$ = 12.0 Hz, $J_{5ax,4}$ = 7.9 Hz, 1H, $H5_{ax}$), 3.03 (s, 1H, OH). Spectrum contains trace amounts of EtOAc. **^{13}C NMR** (101 MHz, $CDCl_3$) δ 167.1 (1C, C_q), 165.2 (1C, C_q) (*C*O, Bz), 133.8 (1C, Ar*C*H), 133.6 (1C, Ar*C*H), 133.0 (1C, Ar*C*), 132.8 (2C, Ar*C*H), 130.2 (2C, Ar*C*H), 130.0 (2C, Ar*C*H), 129.3 (1C, Ar*C*), 129.2 (2C, Ar*C*H), 128.9 (1C, Ar*C*), 128.7 (2C, Ar*C*H), 128.7 (2C, Ar*C*H), 128.3 (1C, Ar*C*H), 86.8 (C1), 76.0 (C3), 70.2 (C2), 68.4 (C4), 67.6 (C5). **HRMS:** calcd. for $C_{25}H_{22}O_6SNa^+$ *m/z* 473.1029; found *m/z* 473.1048.

Phenyl 2,3-di-*O*-levulinoyl-1-thio-β-D-xylopyranoside (92)

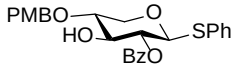


A solution of **90** (1.983 g, 5.47 mmol, 1.0 equiv.) and 4-oxopentanoic acid (1.606 g, 13.83 mmol, 2.5 equiv.) in CH_2Cl_2 (8.5 mL) was cooled to 0 °C. Then a solution of DMAP (0.070 g, 0.58 mmol, 0.11 equiv.) and DDC (3.418 g, 16.81 mmol, 3.1 equiv.) in CH_2Cl_2 (1.0 mL) was added. The solution was stirred at r.t. overnight, after which point it was filtered and conc. *in vacuo*. The residue was dissolved in CH_2Cl_2 (20.0 mL) and H_2O (2.0 mL) and stirred vigorously. DDQ (1.864 g, 8.21 mmol, 1.5 equiv.) was added. After stirring at r.t. for 4.5 h TLC analysis (hexane/EtOAc 1:1) showed the consumption of the starting material and formation of product. The solution was filtered through a big pad of Celite and washed with aq. sat. $NaHCO_3$ and brine, dried over Na_2SO_4, filtered, and conc. *in vacuo*. The residue was purified by flash column chromatography (hexane/EtOAc 1:1) to provide the product as a light yellow crystalline solid (1.776 g, 74% over 2 steps). $\mathbf{R_f}$(EtOAc) 0.58. $[\boldsymbol{\alpha}]_{\boldsymbol{D}}^{\boldsymbol{298\ K}}$ = -23.4 (*c* 1.0, $CHCl_3$). **^{1}H NMR** (400 MHz, $CDCl_3$) δ 7.48 – 7.46 (m, 2H, ArH_{ortho}), 7.33 – 7.29 (m, 3H, ArH_{meta}, ArH_{para}), 5.02 (t, *J* = 8.4 Hz, 1H, H3), 4.94 (t, *J* = 8.7 Hz, 1H, H2), 4.73 (d, $J_{1,2}$ = 8.8 Hz, 1H, H1), 4.19 (dd, $J_{5eq,5ax}$ = 11.7 Hz, $J_{5eq,4}$ =5.1 Hz, 1H, $H5_{eq}$), 3.84 – 3.79 (m, 1H, H4), 3.39 (dd, $J_{5ax,5eq}$ = 11.7 Hz, $J_{5ax,4}$ = 9.5 Hz, 1H, $H5_{ax}$), 3.16 (bs, 1H, OH), 2.89 – 2.51 (m, 8H, $COCH_2CH_2$, CH_2CH_2COO), 2.17 (2 x s, 6H, CH_3CO). **^{13}C NMR** (101 MHz, $CDCl_3$) δ 208.0 (1C, C_q), 206.4 (1C, C_q) (*C*O, Lev), 173.0 (1C, C_q), 171.4 (1C, C_q) (*C*OO, Lev), 132.7 (1C, ArC_{ipso}), 132.6 (2C, ArC_{ortho}H), 129.1 (2C, ArC_{meta}H), 128.2 (1C, ArC_{para}H), 86.8 (C1), 76.5 (C3), 69.7 (C2), 68.6 (C5), 68.5 (C4), 38.4, 37.9 ($COCH_2CH_2$), 29.9, 29.9 (CH_3CO), 28.3, 28.1 (CH_2CH_2COO). **HRMS:** calcd. for $C_{21}H_{26}O_7SNa^+$ *m/z* 461.1241; found *m/z* 461.1242.



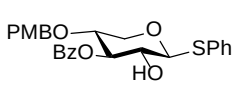
Benzoylation: To a vigorously stirred solution of phenyl 4-*O*-(*p*-methoxy)benzyl-1-thio-β-D-xylopyranoside (**90**) (5.461 g, 15.07 mmol, 1.0 equiv.), tetrabutylammonium hydrogen sulfate (0.736 g, 3.04 mmol, 0.2 equiv.), and BzCl (2.1 mL, 18.08 mmol, 1.2 equiv.) in CH_2Cl_2 (237.0 mL) at -5 °C was added a 1.1 M aq. NaOH solution (40.0 mL, 43.70 mmol, 2.9 equiv.). After 35 min TLC analysis (hexane/EtOAc 2:1) showed the consumption of the starting material and formation of three products. The solution was washed with H_2O, dried over Na_2SO_4, filtered, and conc. *in vacuo*. The residue was purified by flash column chromatography (hexane/EtOAc 8.5:1.5) providing three fractions (**91**: 0.657 g, 8%. **94**: 1.681 g, 24%. **93**: 3.343 g, 48%).

Phenyl 2-*O*-benzoyl-4-*O*-(*p*-methoxy)benzyl-1-thio-β-D-xylopyranoside (93)



$\mathbf{R_f}$(hexane/EtOAc 3:1) 0.35. $[\boldsymbol{\alpha}]_{\boldsymbol{D}}^{\boldsymbol{298\,K}}$ = -13.2 (*c* 1.0, $CHCl_3$). **^{1}H NMR** (400 MHz, $CDCl_3$) δ 8.09 – 8.07 (m, 2H, Ar$H_{\text{ortho(Bz)}}$), 7.61 – 7.57 (m, 1H, Ar$H_{\text{para(Bz)}}$), 7.48 – 7.41 (m, 4H, Ar$H_{\text{meta(Bz)}}$, Ar$H_{\text{ortho(SPh)}}$), 7.29 – 7.24 (m, 5H, Ar$H_{\text{meta(SPh)}}$, Ar$H_{\text{para(SPh)}}$, Ar$H_{\text{ortho(PMB)}}$), 6.89 – 6.86 (m, 2H, Ar$H_{\text{meta(PMB)}}$), 5.04 (dd, $J_{2,1}$ = 9.5 Hz, $J_{2,3}$ = 9.0 Hz, 1H, H2), 4.78 (d, $J_{1,2}$ = 9.6 Hz, 1H, H1), 4.65 (d, J_{gem} = 11.5 Hz, 1H, OCH_2Ph), 4.60 (d, J_{gem} = 11.5 Hz, 1H, OCH_2Ph), 4.10 (dd, $J_{5\text{eq},5\text{ax}}$ = 11.5 Hz, $J_{5\text{eq},4}$ = 5.1 Hz, 1H, H5$_{\text{eq}}$), 3.85 (t, *J* = 8.8 Hz, 1H, H3), 3.81 (s, 3H, OCH_3), 3.61 – 3.55 (m, 1H, H4), 3.30 (dd, $J_{5\text{ax},5\text{eq}}$ = 11.5 Hz, $J_{5\text{ax},4}$ = 10.1 Hz, 1H, H5$_{\text{ax}}$), 2.62 (bs, 1H, OH). **^{13}C NMR** (101 MHz, $CDCl_3$) δ 166.1 (1C, C_{q}, *C*O, Bz), 159.7 (1C, Ar$C_{\text{para(PMB)}}$), 133.5 (1C, ArC_{para}H$_{\text{(Bz)}}$), 132.8 (2C, ArC_{meta}H$_{\text{(Bz)}}$), 132.7 (1C, Ar$C_{\text{ipso(SPh)}}$), 130.1 (2C, ArC_{ortho}H$_{\text{(Bz)}}$), 130.1 (1C, Ar$C_{\text{ipso(Bz)}}$), 129.8 (1C, Ar$C_{\text{ipso(PMB)}}$), 129.7 (2C, ArC_{ortho}H$_{\text{(PMB)}}$), 129.1 (2C, ArC_{ortho}H$_{\text{(SPh)}}$/ArC_{meta}H$_{\text{(SPh)}}$), 128.6 (2C, ArC_{ortho}H$_{\text{(SPh)}}$/ArC_{meta}H$_{\text{(SPh)}}$), 128.2 (1C, ArC_{para}H$_{\text{(SPh)}}$), 114.2 (2C, ArC_{meta}H$_{\text{(PMB)}}$), 86.8 (C1), 77.4 (C4), 76.1 (C3), 73.1 (O*C*H$_2$Ph), 73.0 (C2), 67.6 (C5), 55.4 (O*C*H$_3$). **HRMS:** calcd. for $C_{26}H_{26}O_6SNa^+$ *m/z* 489.1342; found *m/z* 489.1343.

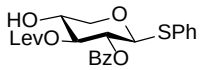
Phenyl 3-*O*-benzoyl-4-*O*-(*p*-methoxy)benzyl-1-thio-β-D-xylopyranoside (94)



$\mathbf{R_f}$(hexane/EtOAc 4:1) 0.37. $[\boldsymbol{\alpha}]_{\boldsymbol{D}}^{\boldsymbol{298\,K}}$ = -125.8 (*c* 1.0, $CHCl_3$). **^{1}H NMR** (400 MHz, $CDCl_3$) δ 8.15 (dd, *J* = 8.4 Hz, *J* = 1.3 Hz, 2H, Ar$H_{\text{ortho(Bz)}}$), 7.63 – 7.58 (m, 1H, Ar$H_{\text{para(Bz)}}$), 7.55 – 7.52 (m, 2H, Ar$H_{\text{ortho(SPh)}}$), 7.50 – 7.46 (m, 2H, Ar$H_{\text{meta(Bz)}}$), 7.35 – 7.29 (m, 3H, Ar$H_{\text{meta(SPh)}}$, Ar$H_{\text{para(SPh)}}$), 7.24 – 7.21 (m, 2H, Ar$H_{\text{ortho(PMB)}}$), 6.84 – 6.80 (m, 2H, Ar$H_{\text{meta(PMB)}}$),

5.35 (t, J = 5.7 Hz, 1H, H3), 5.13 (d, $J_{1,2}$ = 4.9 Hz, 1H, H1), 4.63 (d, J_{gem} = 11.8 Hz, 1H, OCH_2Ph), 4.60 (d, J_{gem} = 11.8 Hz, 1H, OCH_2Ph), 4.41 – 4.37 (m, 1H, H5), 3.84 – 3.81 (m, 1H, H2), 3.78 (s, 3H, OCH_3), 3.70 – 3.62 (m, 2H, H4, H5'), 3.57 (bs, 1H, OH). **^{13}C NMR** (101 MHz, $CDCl_3$) δ 165.9 (1C, C_q, CO, Bz), 159.6 (1C, Ar$C_{para(PMB)}$), 134.2 (1C, Ar$C_{ipso(SPh)}$), 133.6 (1C, ArC_{para}H$_{(Bz)}$), 132.1 (2C, ArC_{ortho}H$_{(SPh)}$), 130.2 (2C, ArC_{ortho}H$_{(Bz)}$), 129.7 (2C, ArC_{ortho}H$_{(PMB)}$), 129.5 (1C, Ar$C_{ipso(Bz)}$/Ar$C_{ipso(PMB)}$), 129.3 (1C, Ar$C_{ipso(Bz)}$/Ar$C_{ipso(PMB)}$), 129.2 (2C, ArC_{meta}H$_{(SPh)}$), 128.6 (2C, ArC_{meta}H$_{(Bz)}$), 127.9 (1C, ArC_{para}H$_{(SPh)}$), 114.1 (2C, ArC_{meta}H$_{(PMB)}$), 89.2 (C1), 73.1 (C4), 72.0 (OCH$_2$Ph), 71.7 (C3), 70.4 (C2), 62.7 (C5), 55.4 (OCH$_3$).

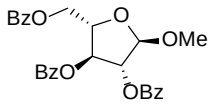
Phenyl 2-*O*-benzoyl-3-*O*-levulinoyl-1-thio-β-D-xylopyranoside (84)



To an ice-cooled solution of **93** (3.943 g, 8.45 mmol, 1.0 equiv.) and 4-oxopentanoic acid (1.223 g, 10.53 mmol, 1.2 equiv.) in CH_2Cl_2 (14.0 mL) was added DMAP (0.104 g, 0.85 mmol, 0.10 equiv.) and DDC (2.611 g, 12.65 mmol, 1.5 equiv.). The solution was left to stir at r.t. After 3.5 h, TLC analysis (hexane/EtOAc 2:1) showed the consumption of the starting material and formation of product. The reaction mixture was filtered and conc. *in vacuo*. The residue was dissolved in CH_2Cl_2 (60.0 mL) and H_2O (6.0 mL) and stirred vigorously. DDQ (2.876 g, 12.67 mmol, 1.5 equiv.) was added. The mixture was left to stir at r.t. overnight in the dark. TLC analysis (hexane/EtOAc 2:1) showed the consumption of the starting material and formation of product. The solution was filtered through a big pad of Celite and washed with aq. sat. $NaHCO_3$ and brine, dried over Na_2SO_4, filtered, and conc. *in vacuo*. The residue was purified by flash column chromatography (hexane/EtOAc 2:1) to provide the product as a white crystalline compound (2.935 g, 78% over 2 steps). $\mathbf{R_f}$(hexane/EtOAc 1:1) 0.28. $[\boldsymbol{\alpha}]_{\boldsymbol{D}}^{\boldsymbol{298\,K}}$ = -1.8 (c 1.0, $CHCl_3$). **^{1}H NMR** (400 MHz, $CDCl_3$) δ 8.03 – 8.01 (m, 2H, Ar$H_{ortho(Bz)}$), 7.59 (t, $J_{14,13}$ = 7.4 Hz, 1H, Ar$H_{para(Bz)}$), 7.47 – 7.44 (m, 4H, Ar$H_{meta(Bz)}$, Ar$H_{ortho(SPh)}$), 7.29 – 7.28 (m, 3H, Ar$H_{meta(SPh)}$, Ar$H_{para(SPh)}$), 5.23 – 5.16 (m, 2H, H2, H3), 4.94 – 4.91 (m, 1H, H1), 4.29 (dd, $J_{5eq,5ax}$ = 11.8 Hz, $J_{5eq,4}$ = 4.9 Hz, 1H, H5$_{eq}$), 3.94 – 3.89 (m, 1H, H4), 3.49 (dd, $J_{5ax,5eq}$ = 11.7 Hz, $J_{5ax,4}$ = 9.1 Hz, 1H, H5$_{ax}$), 3.14 (bs, 1H, OH), 2.81 – 2.65 (m, 2H, COCH_2CH$_2$), 2.58 – 2.51 (m, 1H, CH$_2$CH_2COO), 2.44 – 2.37 (m, 1H, CH$_2$CH_2COO), 2.11 (s, 3H, COCH_3). **^{13}C NMR** (101 MHz, $CDCl_3$) δ 207.7 (1C, C_q, CO, Lev), 172.8 (1C, C_q, COO, Lev), 165.3 (1C, CO, Bz), 133.6 (1C, ArC_{para}H$_{(Bz)}$), 132.9 (1C, Ar$C_{ipso(SPh)}$), 132.7 (2C, ArC_{ortho}H$_{(SPh)}$), 130.0 (2C, ArC_{ortho}H$_{(Bz)}$), 129.4 (1C, Ar$C_{ipso(Bz)}$), 129.1 (2C, ArC_{meta}H$_{(SPh)}$/ArC_{meta}H$_{(Bz)}$), 128.7 (2C, ArC_{meta}H$_{(SPh)}$/ArC_{meta}H$_{(Bz)}$), 128.2 (1C, ArC_{para}H$_{(SPh)}$), 87.0 (C1), 76.4 (C3), 70.3 (C2), 68.5 (C4),

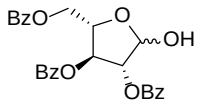
68.3 (C5), 38.5 (CO*C*H$_2$CH$_2$), 29.8 (CO*C*H$_3$), 28.3 (CH$_2$*C*H$_2$COO). **HRMS:** calcd. for $C_{23}H_{24}O_7SNa^+$ *m/z* 467.1135; found *m/z* 467.1136.

2,3,5-tri-*O*-benzoyl-1-*O*-methyl-α-L-arabinofuranose (96)



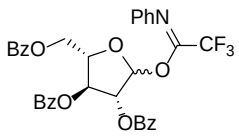
AcCl (5.0 mL, 70.32 mmol, 1.1 equiv.) was slowly added to methanol (60.0 mL) cooled to 0 °C. The methanolic hydrogen chloride solution was then added to a vigorously stirred suspension of L-arabinose (**95**) (10.035 g, 66.84 mmol, 1.0 equiv.) in methanol (200 mL). The reaction mixture was left to stir at r.t. for 4.5 h, at which point TLC analysis (DCM/MeOH 4:1) showed the consumption of the starting material and formation of two products. The reaction mixture was concentrated under co-evaporation with DCM. The resulting residue was cooled to 0 °C and dissolved in pyridine (60.0 mL). BzCl (62.0 mL, 533.72 mmol, 8.0 equiv.) was slowly added over 25 min, and the reaction mixture was left to stir at r.t. overnight. TLC analysis (DCM/MeOH 4:1) showed the consumption of the starting material and TLC analysis (hexane/EtOAc 3:1) showed the formation of product. H_2O (5.0 mL) was added to quench the reaction which was stirred for additional 5 min. Then the mixture was diluted with CH_2Cl_2 (500 mL) and washed with H_2O, 1 M HCl, and $NaHCO_3$, dried over Mg_2SO_4, filtered, and conc. under co-evaporation with toluene. The residue was crystallised with absolute EtOH and addition of pentane was done to promote the crystallisation. The product was obtained as white crystals (14.127 g, 44% over 2 steps). $\mathbf{R_f}$(hexane/EtOAc 2:1) 0.44. **^{1}H NMR** (400 MHz, $CDCl_3$) δ 8.09 – 8.00 (m, 6H, Ar*H*$_{ortho}$), 7.62 – 7.56 (m, 2H), 7.53 – 7.49 (m, 1H) (Ar*H*$_{para}$), 7.46 (t, *J* = 7.7 Hz, 2H), 7.40 (t, *J* = 7.8 Hz, 2H), 7.30 (t, *J* = 7.8 Hz, 2H) (Ar*H*$_{meta}$), 5.59 (d, *J* = 5.2 Hz, 1H, H3), 5.52 (d, *J* = 1.3 Hz, 1H, H2), 5.18 (s, 1H, H1), 4.85 (dd, $J_{5,5'}$ = 11.9 Hz, $J_{5,4}$ = 3.4 Hz, 1H, H5), 4.70 (dd, $J_{5',5}$ = 11.9 Hz, $J_{5',4}$ = 4.8 Hz, 1H, H5), 4.59 – 4.56 (m, 1H, H4), 3.50 (s, 3H, OC*H*$_3$). **^{13}C NMR** (101 MHz, $CDCl_3$) δ 166.4 (1C, *C*$_q$), 166.0 (1C, *C*$_q$), 165.6 (1C, *C*$_q$) (*C*O, Bz), 133.7 (1C, Ar*C*H), 133.6 (1C, Ar*C*H), 133.2 (1C, Ar*C*H) (*C*$_{para}$), 130.1 (2C, Ar*C*H), 130.0 (2C, Ar*C*H), 129.9 (2C, Ar*C*H) (*C*$_{ortho}$), 129.2 (1C, Ar*C*), 129.2 (2C, Ar*C*) (*C*$_{ipso}$), 128.7 (2C, Ar*C*H), 128.6 (2C, Ar*C*H), 28.5 (2C, Ar*C*H) (*C*$_{meta}$), 107.0 (C1), 82.4 (C2), 81.0 (C4), 78.1 (C3), 63.9 (C5), 55.2 (O*C*H$_3$). NMR data are in accordance with literature values.[149]

2,3,5-tri-*O*-benzoyl-α/β-L-arabinofuranose (97)



Furanoside **96** (1.922 g, 4.03 mmol, 1.0 equiv.) was dissolved in AcOH (13.0 mL) and 33% HBr·AcOH (11.0 mL, 60.75 mmol., 15.0 equiv.) was added under an atm. of N_2 and the solution was left to stir at r.t. After 3.5 h TLC analysis (hexane/EtOAc 2:1) showed consumption of the starting material and formation of product. EtOAc was added and the mixture washed with brine and H_2O, dried over $MgSO_4$, filtered, and conc. *in vacuo*. The residue was purified by flash column chromatography (hexane/EtOAc 7:1), and the α/β (ratio α:β 1:0.35) mixture of the product was obtained as a white foam (1.445 g, 77%). $\mathbf{R_f}$(hexane/EtOAc 2:1) 0.41. **^{1}H NMR** (400 MHz, $CDCl_3$) δ 8.13 – 8.03 (m, 8H, Ar*H*), 7.66 – 7.32 (m, 12H, Ar*H*), 5.91 (t, *J* = 5.4 Hz, 0.35H, $H3_\beta$), 5.86 (d, *J* = 4.4 Hz, 0.34H, $H1_\beta$), 5.71 (s, 1H, $H1_\alpha$), 5.61 – 5.57 (m, 2.35H, $H2_\alpha$, $H3_\alpha$, $H2_\beta$), 4.89 – 4.75 (m, 2.70H, $H4_\alpha$, $H5_\alpha$, $H5_\beta$, $H5'_\beta$), 4.69 (dd, *J* = 11.5 Hz, *J* = 5.2 Hz, 1H, $H5'_\alpha$), 4.51 – 4.47 (m, 0.34H, $H4_\beta$). **^{13}C NMR** (101 MHz, $CDCl_3$) δ 166.6 (1C, $C_{q,\beta}$), 166.4 (1C, $C_{q,\alpha}$), 166.1 (1C, $C_{q,\beta}$), 166.0 (1C, $C_{q,\alpha}$), 165.9 (1C, $C_{q,\beta}$), 165.7 (1C, $C_{q,\alpha}$) (*C*O, Bz), 134.0 (Ar*C*H), 133.8 (Ar*C*H), 133.7 (Ar*C*H), 133.7 (Ar*C*H), 133.3 (Ar*C*H), 133.3 (Ar*C*H), 132.2 (Ar*C*H), 130.3 (Ar*C*H), 130.1 (Ar*C*H), 130.1 (Ar*C*H), 130.0 (Ar*C*H), 129.9 (Ar*C*H), 129.9 (Ar*C*H) 129.8 (Ar*C*), 129.1 (Ar*C*), 129.1 (Ar*C*), 128.7 (Ar*C*H), 128.7 (Ar*C*H), 128.6 (Ar*C*H), 128.5 (Ar*C*H), 128.5 (Ar*C*H), 128.4 (Ar*C*H), 101.2 ($C1_\alpha$), 95.7 ($C1_\beta$), 82.6 ($C2_\alpha$), 81.6 ($C4_\alpha$), 79.3 ($C4_\beta$), 78.1 ($C3_\alpha$), 77.8 ($C2_\beta$), 76.7 ($C3_\beta$), 65.9 ($C5_\beta$), 64.1 ($C5_\alpha$). NMR data are in accordance with literature values.[125]

2,3,5-tri-*O*-benzoyl-α/β-L-arabinofuranosyl 1-(*N*-phenyl)-2,2,2-trifluoroacetamidate (98)



Furanose **97** (0.500 g, 1.08 mmol, 1.0 equiv.) was dissolved in acetone (4.0 mL) and Cs_2CO_3 (0.704 g, 2.16 mmol, 2.0 equiv.) and (*Z*)-2,2,2-trifluoro-*N*-phenyl-acetimidoyl chloride (0.35 mL, 2.21 mmol, 2.0 equiv.) were added. The mixture was left to stir at r.t. for 2 h, at which point TLC analysis (hexane/EtOAc 3:1) showed consumption of the starting material and formation of products. The mixture was filtered through a pad of Celite, conc. *in vacuo*., and purified by flash column chromatography (hexane/EtOAc 10:1). The product was obtained as white crystals, where it was possible to separate some of the α-isomer (0.281 g, 41%) from the mixture (α/β ratio 1:2) (0.292 g, 43%) (overall α/β ratio 2:1). $\mathbf{R_f}$(α)(hexane/EtOAc 3:1) 0.57, $\mathbf{R_f}$ (β)(hexane/EtOAc 3:1) 0.52. $[\boldsymbol{\alpha}]_{\boldsymbol{D}}^{\boldsymbol{298\ K}}$(α)= +17.8 (*c* = 1.0, $CHCl_3$). **^{1}H NMR** (400 MHz, $CDCl_3$) δ 8.10 – 8.08 (m, 5H, Ar*H*), 8.06 –

8.04 (m, 3H, Ar*H*), 8.01 (d, *J* = 7.5 Hz, 1H, Ar*H*) 7.66 – 7.38 (m, 11.5H, Ar*H*), 7.32 – 7.26 (m, 3H, Ar$H_{meta(NPh)\alpha}$, Ar*H*), 7.17 (t, *J* = 7.8 Hz, 2H, Ar$H_{meta(NPh)\beta}$), 7.09 (t, *J* = 7.5 Hz, 0.5H, Ar$H_{para(NPh)\alpha}$), 7.03 (t, *J* = 7.5 Hz, 1H, Ar$H_{meta(NPh)\beta}$), 6.88 (d, *J* = 7.8 Hz, 1H, Ar$H_{ortho(NPh)\alpha}$), 6.83 (bs, 1H, H1$_\beta$), 6.58 (d, *J* = 7.7 Hz, 2H, Ar$H_{ortho(NPh)\beta}$), 6.02 – 5.99 (m, 1H, H3$_\beta$), 5.93 – 5.91 (m, 1H, H2$_\beta$), 5.80 (s, 0.5H, H2$_\alpha$), 5.67 – 5.67 (m, 0.5H, H3$_\alpha$), 4.84 – 4.80 (m, 2H, H4$_\alpha$, H5$_\alpha$, H5$_\beta$), 4.76 – 4.71 (m, 1.5H, H5'$_\alpha$, H5'$_\beta$), 4.67 – 4.63 (m, 1H, H4$_\beta$), 1.56 (s, H_2O). **^{13}C NMR** (101 MHz, $CDCl_3$) δ 166.3 (1C, $C_{q,\alpha}$), 166.2 (1C, $C_{q,\beta}$), 165.9 (1C, $C_{q,\beta}$), 165.7 (1C, $C_{q,\alpha}$), 165.6 (1C, $C_{q,\beta}$), 165.2 (1C, $C_{q,\alpha}$) (*C*O, Bz), 143.5 (1C, C_q), 143.4 (1C, C_q) (O*C*=NPh$(CF_3)_\alpha$, O*C*=NPh$(CF_3)_\beta$), 134.0 (1C, Ar*C*H), 134.0 (2C, Ar*C*H), 133.9 (2C, Ar*C*H), 133.3 (2C, Ar*C*H), 130.1 (4C, Ar*C*H), 130.1 (4C, Ar*C*H), 130.0 (2C, Ar*C*H), 129.9 (4C, Ar*C*H), 129.7 (1C, Ar*C*), 129.7 (1C, Ar*C*), 129.0 (1C, Ar*C*), 128.9 (2C, Ar*C*H), 128.8 (4C, Ar*C*H), 128.7, 128.5 (4C, Ar*C*H), 124.6 (1C, Ar$C_{para}H_{(NPh)\alpha}$), 124.3 (1C, Ar$C_{para}H_{(NPh)\beta}$), 119.7 (2C, Ar$C_{ortho}H_{(NPh)\alpha}$), 119.3 (2C, Ar$C_{ortho}H_{(NPh)\beta}$), 102.4 (C1$_\alpha$), 97.0 (C1$_\beta$), 84.4 (C4$_\alpha$), 80.9 (C2$_\alpha$), 80.7 (C4$_\beta$), 77.4 (C3$_\alpha$), 76.3 (C2$_\beta$), 75.7 (C3$_\beta$), 65.0 (C5$_\beta$), 63.7 (C5$_\alpha$). **HRMS:** calcd. for $C_{34}H_{26}F_3NO_8Na^+$ *m/z* 656.1503; found *m/z* 656.1499. NMR data for the α/β mixture are in accordance with literature values.[124,125]

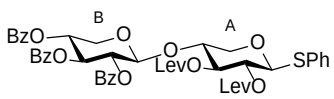
A: General procedure for glycosylation with a thiophenyl donor: Crushed molecular sieves (3 Å) is added a 2-necks Schlenk flask, where the middle neck is fitted with a glass stopper, and the other with a septum. The flask is placed under vacuum and heated. Afterwards an atm. of N_2 is applied. The donor (1.0 equiv.) dissolved in dry CH_2Cl_2 (c = 0.12 M) is added to the flask together with AgOTf (2.0 equiv.) dissolved in dry toluene (c = 0.36 M). Stirring of the reaction mixture is initiated and the glass stopper is exchanged for a thermometer under N_2 pressure, and the solution is cooled to -65 °C. 4-Nitrobenzene-sulfenyl chloride (1.0 equiv.) is dissolved in dry CH_2Cl_2 (c = 0.72 M) and slowly added as to not raise the temperature above -60 °C. The mixture is left to stir for approximately 10 min or until activation is complete. The acceptor (0.9 equiv.) is dissolved in dry CH_2Cl_2 (c = 0.65 M) and added quickly. The temperature is kept between -55 °C and -50 °C until TLC analysis shows completion of the reaction. Afterwards the mixture is allowed to heat to -15 °C over 10 min at which point Et_3N (3.0 equiv.) is added. The solution is filtered through a pad of Celite and conc. *in vacuo*. The residue is then purified by flash column chromatography to afford the product.

B: General procedure for deprotection of levulinoyl groups: The linear xylan saccharide (1.0 equiv.) is dissolved in AcOH (c = 0.40 M) and pyridine (c = 0.20 M). A mixture of a 50% solution of hydrazine hydrate (10.0 equiv.) dissolved in AcOH (c = 20 M) and pyridine (c = 10 M), is added. The reaction mixture is left to stir at r.t. until TLC shows consumption of the starting material and formation of a product. The reaction is stopped by addition of acetone (300 equiv.) and left to stir at r.t for 30 min. EtOAc is added, and the mixture is washed with 1 M HCl, $NaHCO_3$, and H_2O, dried over Na_2SO_4, filtered, and conc. *in vacuo*. The product is obtained in sufficient purity without further purification.

C: General procedure for glycosylation with 1-(*N*-phenyl)-2,2,2-trifluoroacetimidate: The deprotected xylan-chain (1.0 equiv.) and 2,3,5-tri-*O*-benzoyl-α/β-L-arabinofuranosyl 1-(*N*-phenyl)-2,2,2-trifluoroacetamidate (**98**) (4.0 equiv.) are co-evaporated with toluene twice and left under high vacuum overnight. The mixture is dissolved in dry CH_2Cl_2 (20 mL/0.1 g deprotected xylan-chain) and cooled to -40 °C. TMSOTf (0.1 equiv.) is added from a freshly made stock solution of TMSOTf in dry CH_2Cl_2. The reaction mixture is kept at -40 °C for an appropriate reaction time after which Et_3N (0.7 equiv.) is added, and the mixture is conc. *in vacuo*. The residue is purified by flash column chromatography to afford the desired product.

D: General procedure for deprotection of anomeric thiophenyl group: The arabinoxylan saccharide (1.0 equiv.) is dissolved in acetone/H_2O 9:1 (15 mL/mmol), NBS (4.0 equiv.) is added and the mixture is left to stir at r.t. until TLC analysis shows completion of the reaction. EtOAc is added to the mixture, which is then washed with $NaHCO_3$, and H_2O, dried over Na_2SO_4, filtered, and conc. *in vacuo*. The residue is purified by flash column chromatography.

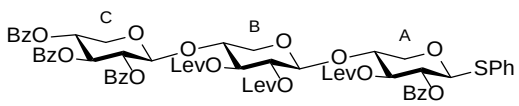
Phenyl 2,3,4-tri-*O*-benzoyl-β-D-xylopyranosyl-(1→4)-2,3-di-*O*-levulinoyl-1-thio-β-D-xylopyranoside (105)



General procedure A: Crushed molecular sieves (1.984 g), donor **80** (0.402 g, 0.72 mmol, 1.0 equiv.), AgOTf (0.378 g, 1.47 mmol, 2.0 equiv.), 4-nitrobenzenesulfenyl chloride (0.141 g, 0.74 mmol, 1.0 equiv.), and acceptor **92** (0.286 g, 0.65 mmol, 0.9 equiv.). Reaction time 25 min. Et_3N (0.30 mL, 2.15 mmol, 3.0 equiv.). Eluent for flash column chromatography (pentane/EtOAc 3:2). The product was obtained as

white crystals (0.547 g, 95%). **R$_f$**(hexane/EtOAc 1:1) 0.42. $[\boldsymbol{\alpha}]_{\boldsymbol{D}}^{\boldsymbol{298\,K}}$ = -28.1 (*c* 0.5, $CHCl_3$). **^{1}H NMR** (400 MHz, $CDCl_3$) δ 8.02 (d, *J* = 7.2 Hz, 2H, Ar*H*), 7.94 (d, *J* = 7.4 Hz, 4H, Ar*H*), 7.59 – 7.49 (m, 3H, Ar*H*), 7.45 – 7.28 (m, 11H, Ar*H*), 5.69 (t, *J* = 6.4 Hz, 1H, H3^B), 5.30 (td, *J* = 6.1 Hz, $J_{4B,5Beq}$ = 4.0 Hz, 1H, H4^B), 5.25 (d, *J* = 8.8 Hz, 1H, H3^A) 5.22 – 5.19 (m, 1H, H2^B), 4.92 – 4.87 (m, 2H, H2^A, H1^B), 4.66 (d, $J_{A1,A2}$ = 9.3 Hz, 1H, H1^A), 4.49 (dd, $J_{5Beq,5Bax}$ = 12.4 Hz, $J_{5Beq,4B}$ = 3.8 Hz, 1H, H5$^B{}_{eq}$), 4.09 (dd, $J_{5Aeq,5Aax}$ = 11.9 Hz, $J_{5Aeq,4A}$ = 5.2 Hz, 1H, H5$^A{}_{eq}$), 3.90 (td, *J* = 9.4 Hz, $J_{4A,5Aeq}$ = 5.2 Hz, 1H, H4^A), 3.75 (dd, $J_{5Bax,5Beq}$ =12.4 Hz, $J_{5Bax,4B}$ = 6.0 Hz, 1H, H5$^B{}_{ax}$), 3.32 (dd, $J_{5Aax,5Aeq}$ = 11.7 Hz, $J_{5Aeq,4A}$ = 9.9 Hz, 1H, H5$^A{}_{ax}$), 2.88 – 2.53 (m, 8H, COC*H*$_2$CH$_2$, CH$_2$C*H*$_2$COO), 2.19 (s, 3H, C*H*$_3$CO), 2.08 (s, 3H, C*H*$_3$CO). **^{13}C NMR** (101 MHz, $CDCl_3$) δ 206.5 (1C, C_q), 206.4 (1C, C_q) (*C*O, Lev), 171.9 (1C, C_q), 171.5 (1C, C_q) (*C*OO, Lev), 165.7 (1C, C_q), 165.4 (1C, C_q), 165.2 (1C, C_q) (*C*O, Bz), 133.6 (1C, Ar*C*H), 133.5 (1C, Ar*C*H), 133.5 (1C, Ar*C*H), 132.9 (2C, Ar*C*H), 132.2 (1C, Ar*C*), 130.1 (2C, Ar*C*H), 130.1 (2C, Ar*C*H), 129.9 (2C, Ar*C*H), 129.4 (1C, Ar*C*), 129.2 (1C, Ar*C*), 129.1 (1C, Ar*C*), 129.1 (2C, Ar*C*H), 128.6 (4C, Ar*C*H), 128.3 (1C, Ar*C*H), 99.4 (C1^B), 86.6 (C1^A), 74.9 (C4^A), 73.5 (C3^A), 70.3 (C2^A), 70.2 (C2^B), 69.6 (C3^B), 68.7 (C4^B), 66.7 (C5^A), 61.1 (C5^B), 37.9, 37.8 (CO*C*H$_2$CH$_2$), 30.0, 29.8 (*C*H$_3$CO), 28.1, 28.0 (CH$_2$*C*H$_2$COO). **HRMS:** calcd. for $C_{21}H_{26}O_7SNa^+$ *m/z* 9005.2449; found *m/z* 905.2444.

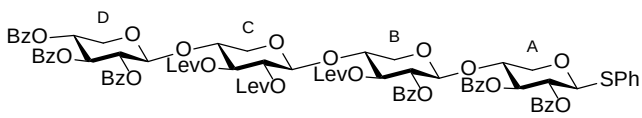
Phenyl 2,3,4-tri-*O*-benzoyl-β-D-xylopyranosyl-(1→4)-2,3-di-*O*-levulinoyl-β-D-xylopyranosyl-(1→4)-2-*O*-benzoyl-3-*O*-levulinoyl-1-thio-β-D-xylopyranoside (106)



General procedure A: Crushed molecular sieves (1.229 g), donor **105** (0.252 g, 0.29 mmol, 1.0 equiv.), AgOTf (0.148 g, 0.58 mmol, 2.0 equiv.), 4-nitrobenzenesulfenyl chloride (0.054 g, 0.28 mmol, 1.0 equiv.), and acceptor **84** (0.114 g, 0.26 mmol, 0.9 equiv.). Reaction time 25 min. Et_3N (0.12 mL, 0.86 mmol, 3.0 equiv. Eluent for flash column chromatography (heptane/toluene/EtOAc 2:1:1). The product was obtained as light yellow crystals (0.162 g, 52%). **R$_f$**(hexane/EtOAc 1:1) 0.35. $[\boldsymbol{\alpha}]_{\boldsymbol{D}}^{\boldsymbol{298\,K}}$ = -53.8 (*c* 1.0, $CHCl_3$). **^{1}H NMR** (400 MHz, $CDCl_3$) δ 8.03 – 7.99 (m, 6H, Ar*H*), 7.96 – 7.93 (m, 4H, Ar*H*), 7.60 – 7.50 (m, 5H, Ar*H*), 7.47 – 7.32 (m, 12H, Ar*H*), 7.28 – 7.26 (m, 3H, Ar*H*), 5.69 (t, *J* = 6.6 Hz, 1H, H3^C) 5.30 – 5.25 (m, 2H, H3^A, H4^C), 5.22 (dd, $J_{2C,3C}$ = 6.5 Hz, $J_{2C,1C}$ = 4.8 Hz, 1H, H2^C), 5.17 – 5.10 (m, 2H, H2^A, H3^B), 4.93 (d, $J_{1A,2A}$ = 8.1 Hz, 1H, H1^A), 4.87 (d, $J_{1C,2C}$ = 4.7 Hz, 1H, H1^C), 4.77 (dd, *J* = 8.9 Hz, *J* = 7.0 Hz, 1H, H2^B), 4.48 – 4.44 (m, 2H, H1^B, H5^C), 4.22 (dd, $J_{5Aeq,5Aax}$ = 12.0 Hz, $J_{5Aeq,4A}$ = 4.7 Hz, 1H, H5$^A{}_{eq}$), 3.95 (dd, $J_{5Beq,5Bax}$ = 12.0 Hz, $J_{5Beq,4B}$ = 5.0 Hz, 1H, H5$^B{}_{eq}$), 3.86 – 3.76 (m, 2H, H4^A, H4^B), 3.73 (dd,

$J_{5C',5C}$ = 12.4 Hz, $J_{5C',4C}$ = 6.2 Hz, 1H, H5$^{C'}$), 3.50 (dd, $J_{5Aax,5Aeq}$ = 12.0 Hz, $J_{5ax,4A}$ = 8.6 Hz, 1H, H5$^{A}_{ax}$), 3.27 (dd, $J_{5Bax,5Beq}$ = 12.0 Hz, $J_{5Bax,4B}$ = 9.1 Hz, 1H, H5$^{B}_{ax}$), 2.75 – 2.69 (m, 4H, COCH_2CH$_2$, CH$_2$CH_2COO), 2.62 – 2.36 (m, 8H, COCH_2CH$_2$, CH$_2$CH_2COO), 2.17 (s, 3H), 2.09 (s, 3H), 2.04 (s, 3H) (CH_3CO). **^{13}C NMR** (101 MHz, CDCl$_3$) δ 206.5 (1C, C_q), 206.4 (1C, C_q), 206.1 (1C, C_q) (*C*O, Lev), 172.0 (1C, C_q), 171.5 (1C, C_q), 171.3 (1C, C_q) (*C*OO, Lev), 165.7 (1C, C_q), 165.4 (1C, C_q), 165.3 (1C, C_q), 165.2 (1C, C_q) (*C*O, Bz), 133.5 (1C, Ar*C*H), 133.5 (1C, Ar*C*H), 133.5 (1C, Ar*C*H), 133.4 (1C, Ar*C*H), 132.8 (1C, Ar*C*), 132.7 (2C, Ar*C*H), 130.1 (2C, Ar*C*H), 130.1 (2C, Ar*C*H), 130.0 (2C, Ar*C*H), 129.9 (2C, Ar*C*H), 129.5 (1C, Ar*C*), 129.4 (1C, Ar*C*), 129.2 (1C, Ar*C*), 129.1 (1C, Ar*C*), 129.1 (2C, Ar*C*H), 128.6 (4C, Ar*C*H), 128.1 (1C, Ar*C*H), 100.8 (C1^{B}), 99.3 (C1^{C}), 86.5 (C1^{A}), 75.0 (C4^{A}), 74.7 (C4^{B}), 72.7 (C3^{A}), 72.2 (C3^{B}), 71.4 (C2^{B}), 70.4 (C2^{A}), 70.2 (C2^{C}), 69.7 (C3^{C}), 68.8 (C4^{C}), 65.8 (C5^{A}), 62.9 (C5^{B}), 61.1 (C5^{C}), 37.9, 37.8, 37.8 (CO*C*H$_2$CH$_2$), 29.9, 29.8, 29.8 (*C*H$_3$CO), 28.1, 28.0, 27.9 (CH$_2$*C*H$_2$COO). **HRMS:** calcd. for C$_{64}$H$_{64}$O$_{22}$SNa$^+$ *m/z* 1239.3502; found *m/z* 1239.3483.

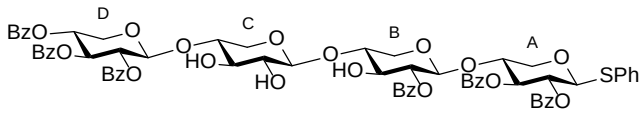
Phenyl 2,3,4-tri-*O*-benzoyl-β-D-xylopyranosyl-(1→4)-2,3-di-*O*-levulinoyl-β-D-xylopyranosyl-(1→4)-2-*O*-benzoyl-3-*O*-levulinoyl-β-D-xylopyranosyl-(1→4)-2,3-di-*O*-benzoyl-1-thio-β-D-xylopyranoside (107)



General procedure A: Crushed molecular sieves (1.219 g), donor **106** (0.414 g, 0.34 mmol, 1.0 equiv.), AgOTf (0.175 g, 0.68 mmol, 2.0 equiv.), 4-nitrobenzenesulfenyl chloride (0.064 g, 0.34 mmol, 1.0 equiv.), and acceptor **81** (0.139 g, 0.31 mmol, 0.9 equiv.). Reaction time 65 min. Et$_3$N (0.14 mL, 1.02 mmol, 3.0 equiv.). Eluent for flash column chromatography (hexane/toluene/acetone 3:2:1). The product was obtained as white crystals (0.358 g, 75%), furthermore the donor **105** was reisolated (0.064g, 15%). **R$_f$**(hexane/EtOAc 1:1) 0.32. $[\boldsymbol{\alpha}]_{\boldsymbol{D}}^{\boldsymbol{298\ K}}$ = -35.1 (*c* 1.0, CHCl$_3$). **^{1}H NMR** (400 MHz, CDCl$_3$) δ 8.01 – 7.92 (m, 12H, Ar*H*), 7.58 – 7.50 (m, 6H, Ar*H*), 7.45 – 7.31 (m, 14H, Ar*H*), 7.28 – 7.24 (m, 3H, Ar*H*), 5.69 (t, *J* = 6.6 Hz, 1H, H3^{D}), 5.60 (t, *J* = 8.0 Hz, 1H, H3^{A}), 5.32 – 5.26 (m, 2H, H2^{A}, H4^{D}), 5.21 (dd, $J_{2D,3D}$ = 6.5 Hz, $J_{2D,1D}$ = 4.8 Hz, 1H, H2^{D}), 5.13 – 5.07 (m, 2H, H3^{B}, H3^{C}), 4.98 – 4.95 (m, 2H, H1^{A}, H2^{B}), 4.84 (d, $J_{1D,2D}$ = 4.7 Hz, 1H, H1^{D}), 4.73 – 4.68 (m, 2H, H1^{B}, H2^{C}), 4.46 (dd, $J_{5Deq,5Dax}$ = 12.4 Hz, $J_{5Deq,4D}$ = 3.9 Hz, 1H, H5$^{D}_{eq}$), 4.21 – 4.15 (m, 2H, H5^{A}, H1^{C}), 4.02 – 3.96 (m, 1H, H4^{A}), 3.88 (dd, $J_{5C,5'C}$ = 11.9 Hz, $J_{5C,4C}$ = 5.1 Hz, 1H, H5^{C}), 3.83 – 3.77 (m, 1H, H4^{C}), 3.72 (dd, $J_{5Dax,5Dax}$ = 12.4 Hz, $J_{5Dax,4D}$ = 6.2 Hz, 1H, H5$^{D}_{ax}$), 3.56 (dd, $J_{5Beq,5Bax}$ = 11.9 Hz, $J_{5Beq,4B}$ = 4.6 Hz, 1H, H5$^{B}_{eq}$), 3.52 – 3.44 (m, 2H, H5^{A}', H4^{B}), 3.20 – 3.14 (m, 2H, H5'^{C}, H5$^{B}_{ax}$), 2.73 – 2.69 (m, 4H, COCH_2CH$_2$,

CH_2CH_2COO), 2.60 – 2.31 (m, 8H, $COCH_2CH_2$, CH_2CH_2COO), 2.17 (s, 3H), 2.08 (s, 3H), 2.03 (s, 3H) (CH_3CO). **^{13}C NMR** (101 MHz, $CDCl_3$) δ 206.4 (1C, C_q), 206.3 (1C, C_q), 206.1 (1C, C_q) (*C*O, Lev), 172.0 (1C, C_q), 171.6 (1C, C_q), 171.3 (1C, C_q) (*C*OO, Lev), 165.6 (1C, C_q), 165.4 (1C, C_q), 165.4 (1C, C_q), 165.3 (1C, C_q), 165.1 (1C, C_q), 165.1 (1C, C_q) (*C*O, Bz), 133.6 (1C, Ar*C*H), 133.5 (1C, Ar*C*H), 133.5 (1C, Ar*C*H), 133.4 (1C, Ar*C*H), 133.4 (1C, Ar*C*H), 133.3 (1C, Ar*C*H), 132.8 (1C, Ar*C*), 132.6 (2C, Ar*C*H), 130.1 (2C, Ar*C*H), 130.0 (2C, Ar*C*H), 130.0 (2C, Ar*C*H), 129.9 (2C, Ar*C*H), 129.9 (2C, Ar*C*H), 129.9 (2C, Ar*C*H), 129.6 (1C, Ar*C*), 129.5 (1C, Ar*C*), 129.4 (1C, Ar*C*), 129.4 (1C, Ar*C*), 129.2 (1C, Ar*C*), 129.1 (1C, Ar*C*), 129.1 (2C, Ar*C*H), 128.6 (2C, Ar*C*H), 128.6 (4C, Ar*C*H), 128.5 (2C, Ar*C*H), 128.1 (1C, Ar*C*H), 100.5 ($C1^C$, $J_{C\text{-}H}$ = 161 Hz), 100.3 ($C1^B$, $J_{C\text{-}H}$ = 164 Hz), 99.3 ($C1^D$, $J_{C\text{-}H}$ = 165 Hz), 86.7 ($C1^A$, $J_{C\text{-}H}$ = 159 Hz), 75.0 ($C4^A$), 74.7 ($C4^C$), 74.5 ($C4^B$), 73.0 ($C3^A$), 72.2 ($C3^C$), 71.4 ($C3^B$), 71.2 ($C2^C$), 71.0 ($C2^B$), 70.4, 70.2 ($C2^D$, $C4^D$), 69.7 ($C3^D$), 68.8 ($C2^A$), 65.8 ($C5^A$), 62.9 ($C5^C$), 62.0 ($C5^B$), 61.1 ($C5^D$), 37.8, 37.8 (2C) ($CO*C*H_2CH_2$), 29.9, 29.8, 29.8 ($*C*H_3CO$), 28.0, 27.9, 27.8 (CH_2*C*H_2COO). **HRMS:** calcd. for $C_{83}H_{80}O_{28}SNa^+$ *m/z* 1579.4449; found *m/z* 1579.4425.

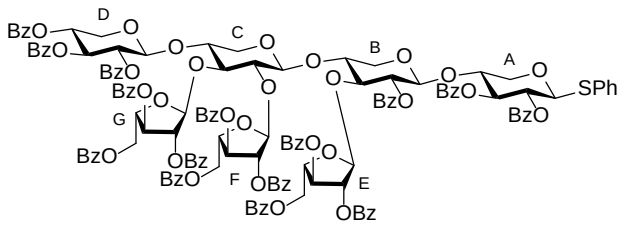
Phenyl 2,3,4-tri-*O*-benzoyl-β-D-xylopyranosyl-(1→4)-β-D-xylopyranosyl-(1→4)-2-*O*-benzoyl-β-D-xylopyranosyl-(1→4)-2,3-di-*O*-benzoyl-1-thio-β-D-xylopyranoside (108)



General procedure B: Tetrasaccharide **106** (0.252 g, 0.16 mmol, 1.0 equiv.), and 50% solution of hydrazine hydrate (0.10 mL, 1.59 mmol, 10.0 equiv.). Reaction time 10 min. Eluent for TLC (hexane/toluene/acetone 2:1:1). Acetone (3.5 mL, 48.0 mmol, 300 equiv.). The product was a white powder (0.202 g, 99%). **R_f**(hexane/acetone 1:1) 0.64. $[\alpha]_D^{298\,K}$ = -9.7 (*c* 1.0, $CHCl_3$). **^{1}H NMR** (400 MHz, $CDCl_3$) δ 8.03 – 7.91 (m, 12H, Ar*H*), 7.59 – 7.33 (m, 20H, Ar*H*), 7.28 – 7.25 (m, 3H, Ar*H*), 5.80 (t, *J* = 8.5 Hz, 1H, $H3^D$), 5.58 (t, $J_{A3,A2}$ = 7.7 Hz, 1H, $H3^A$), 5.40 – 5.34 (m, 2H, $H2^D$, $H4^D$), 5.31 (t, *J* = 7.8 Hz, 1H, $H2^A$), 4.99 (d, $J_{1A,2A}$ = 7.8 Hz, 1H, $H1^A$), 4.92 (dd, $J_{2B,3B}$ = 8.6 Hz, $J_{2B,1B}$ = 7.0 Hz, 1H, $H2^B$), 4.79 (d, $J_{1D,2D}$ = 6.6 Hz, 1H, $H1^D$), 4.63 (d, $J_{1B,2B}$ = 6.9 Hz, 1H, $H1^B$), 4.50 (dd, $J_{5Deq,5Dax}$ = 11.9 Hz, $J_{5Deq,4D}$ = 4.9 Hz, 1H, $H5^D{}_{eq}$), 4.18 (dd, $J_{5Aeq,5Aax}$ = 12.1 Hz, $J_{5Aeq,4A}$ = 4.6 Hz, 1H, $H5^A{}_{eq}$), 4.14 (d, $J_{1C,2C}$ = 7.4 Hz, 1H, $H1^C$), 4.01 – 3.95 (m, 1H, $H4^A$), 3.75 – 3.42 (m, 9H, $H5^A{}_{ax}$, $H3^B$, $H4^B$, $H5^B$, $H3^C$, $H4^C$, $H5^C$, $H5^D{}_{ax}$), 3.37 – 3.33 (m, 1H, $H2^C$), 3.12 (m, 2H, $H5^{B'}$, $H5^{C'}$). **^{13}C NMR** (101 MHz, $CDCl_3$) δ 165.7 (1C, C_q), 165.6 (1C, C_q), 165.6 (1C, C_q), 165.5 (1C, C_q), 165.3 (1C, C_q), 165.2 (1C, C_q) (CO, Bz), 133.7 (1C, Ar*C*H), 133.6 (1C, Ar*C*H), 133.5 (1C, Ar*C*H), 133.4 (1C, Ar*C*H), 133.3 (1C, Ar*C*H), 133.3 (1C, Ar*C*H), 132.9 (C_q, Ar*C*), 132.5 (2C, Ar*C*H), 130.0 (2C, Ar*C*H), 130.0 (2C, Ar*C*H), 129.9 (2C, Ar*C*H),

129.9 (3C, Ar*C*H), 129.8 (2C, Ar*C*H), 129.6 (1C, Ar*C*), 129.6 (1C, Ar*C*), 129.4 (1C, Ar*C*), 129.1 (2C, Ar*C*H), 129.0 (1C, Ar*C*), 129.0 (1C, Ar*C*), 128.9 (1C, Ar*C*), 128.7 (2C, Ar*C*H), 128.6 (3C, Ar*C*H), 128.5 (2C, Ar*C*H), 128.5 (3C, Ar*C*H), 128.1 (1C, Ar*C*H), 102.5 ($C1^{C}$), 101.4 ($C1^{B}$), 101.0 ($C1^{D}$), 86.6 ($C1^{A}$), 78.8 ($C4^{C}$), 77.4 ($C4^{B}$), 75.7 ($C4^{A}$), 74.0 ($C3^{C}$), 73.7 ($C2^{B}$), 72.8 , 72.7 ($C3^{A}$, $C3^{B}$), 72.1 ($C2^{C}$), 71.1, 70.9 ($C2^{D}$, $C3^{D}$), 70.3 ($C2^{A}$), 69.3 ($C4^{D}$), 65.7 ($C5^{A}$), 63.1 ($C5^{B}$/$C5^{C}$), 63.0 ($C5^{B}$/$C5^{C}$), 62.5 ($C5^{D}$). **HRMS:** calcd. for $C_{68}H_{62}O_{22}SNa^{+}$ *m/z* 1285.3345; found *m/z* 1285.3335.

Phenyl 2,3,4-tri-*O*-benzoyl-β-D-xylopyranosyl-(1→4)-[(1→2)-2,3,5-tri-*O*-benzoyl-α-L-arabinofuranosyl]- [(1→3)-2,3,5-tri-*O*-benzoyl-α-L-arabinofuranosyl]-β-D-xylopyranosyl-(1→4)-[(1→3)-2,3,5-tri-*O*-benzoyl-α-L-arabinofuranosyl]-2-*O*-benzoyl-β-D-xylopyranosyl-(1→4)-2,3-di-*O*-benzoyl-1-thio-β-D-xylopyranoside (109)



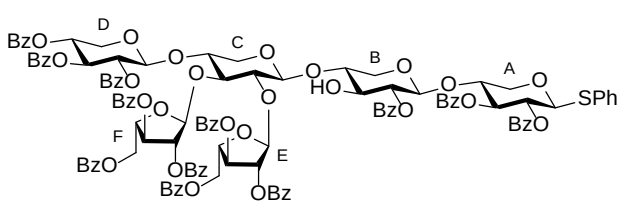
General procedure C: Acceptor **108** (0.088 g, 0.069 mmol, 1.0 equiv.), donor **98** (0.175 g, 0.28 mmol, 4.0 equiv.), and TMSOTf in dry CH_2Cl_2 (0.17 mL, c = 0.04 mmol/mL, 0.1 equiv.). Reaction time 105 min. Et_3N (0.02 mL, 0.14 mmol, 2.0 equiv.). Eluent for flash column chromatography (hexane/toluene/acetone 3:3:1). The product **109** was obtained as white crystals (0.168 g, 91%). $\mathbf{R_f}$(hexane/toluene/acetone 3:3:1) 0.26. $[\boldsymbol{\alpha}]_{\boldsymbol{D}}^{\boldsymbol{298\,K}}$ = -58.4 (*c* 1.0, $CHCl_3$). **^{1}H NMR** (400 MHz, $CDCl_3$) δ 8.22 (d, *J* = 8.1 Hz, 2H, Ar*H*), 8.09 (dd, *J* = 12.7 Hz, *J* = 7.4 Hz, 6H, Ar*H*), 8.03 – 7.93 (m, 14H, Ar*H*), 7.86 – 7.79 (m, 8H, Ar*H*), 7.73 (d, *J* = 7.3 Hz, 2H, Ar*H*), 7.58 (d, *J* = 7.4 Hz, 1H, Ar*H*), 7.62 – 7.10 (m, 46H, Ar*H*), 7.03 (t, *J* = 7.8 Hz, 1H, Ar*H*), 5.69 (s, 1H, $H1^{F}$), 5.63 (t, *J* = 9.6 Hz, 1H, $H4^{D}$), 5.58 (s, 1H, $H1^{G}$), 5.53 (s, 1H), 5.51 (d, *J* = 4.4 Hz, 1H), 5.40 – 5.39 (m, 3H, $H3^{A}$), 5.37 – 5.34 (m, 2H), 5.21 – 5.16 (m, 4H, $H2^{A}$, $H2^{D}$, $H1^{E}$), 5.13 – 5.03 (m, 3H, $H2^{B}$, $H4^{D}$), 5.00 – 4.93 (m, 2H), 4.86 – 4.74 (m, 7H, $H1^{A}$), 4.31 (d, $J_{B1,B2}$ = 6.6 Hz, 1H, $H1^{B}$), 4.02 – 3.96 (m, 2H, $H5^{A}$, $H1^{D}$), 3.90 (t, *J* = 8.7 Hz, 1H, $H3^{B}$), 3.83 – 3.78 (m, 2H, $H1^{C}$, $H3^{C}$), 3.58 – 3.52 (m, 2H, $H4^{B}$, $H2^{C}$), 3.45 (td, *J* = 9.5 Hz, *J* = 5.0 Hz, 1H, $H4^{C}$), 3.41 – 3.35 (m, 2H, $H4^{A}$, $H5^{D}$), 3.32 – 3.24 (m, 2H, $H5^{A'}$, $H5^{B}{}_{eq}$), 3.09 – 3.00 (m, 2H, $H5^{C}$, $H5^{D'}$), 2.92 (dd, $J_{5Bax,5Beq}$ = 12.2 Hz, $J_{5Bax,4B}$ = 7.5 Hz, 1H, $H5^{B}{}_{ax}$), 2.45 (t, *J* = 10.8 Hz, 1H, $H5^{C'}$). **^{13}C NMR** (101 MHz, $CDCl_3$) δ 166.6 (2C, C_q), 166.4 (1C, C_q), 166.3 (1C, C_q), 165.9 (1C, C_q), 165.8 (1C, C_q), 165.8 (1C, C_q), 165.6 (1C, C_q), 165.5 (1C, C_q), 165.3 (1C, C_q), 165.2 (1C, C_q), 164.9 (1C, C_q), 164.9 (1C, C_q), 164.9 (1C, C_q), 164.4 (1C, C_q) (*C*O, Bz), 134.1 (1C, Ar*C*H), 134.0 (1C, Ar*C*H), 133.5 (1C, Ar*C*H), 133.4 (3C, Ar*C*H), 133.4 (1C, Ar*C*H) 133.3 (1C, Ar*C*H), 133.3 (1C, Ar*C*H), 133.1 (1C, Ar*C*H), 133.0

(1C, Ar*C*H), 132.9 (1C, Ar*C*H), 132.8 (1C, Ar*C*H), 132.1 (1C, Ar*C*), 130.3 (2C, Ar*C*H), 130.2 (2C, Ar*C*H), 130.1 (1C, Ar*C*H), 130.1 (3C, Ar*C*H), 130.0 (3C, Ar*C*H), 130.0 (2C, Ar*C*H), 130.0 (3C, Ar*C*H), 129.9 (4C, Ar*C*H), 129.9 (3C, Ar*C*H), 129.9 (4C, Ar*C*H), 129.8 (4C, Ar*C*H), 129.8 (2C, Ar*C*H), 129.7 (2C, Ar*C*H), 129.7 (1C, Ar*C*), 129.6 (1C, Ar*C*), 129.5 (1C, Ar*C*), 129.3 (2C, Ar*C*), 129.3 (1C, Ar*C*), 129.2 (1C, Ar*C*), 129.1 (1C, Ar*C*), 129.0 (2C, Ar*C*H), 129.0 (1C, Ar*C*), 129.0 (1C, Ar*C*), 128.9 (2C, Ar*C*H), 128.6 (2C, Ar*C*H), 128.6 (3C, Ar*C*H), 128.5 (3C, Ar*C*H), 128.5 (2C, Ar*C*H), 128.4 (2C, Ar*C*H), 128.3 (4C, Ar*C*H), 128.3 (2C, Ar*C*H), 128.2 (4C, Ar*C*H), 128.2 (2C, Ar*C*H), 106.1 ($C1^{G}$), 105.7 ($C1^{F}$), 105.4 ($C1^{E}$), 100.1 ($C1^{B}$), 99.8 ($C1^{D}$), 99.6 ($C1^{C}$), 86.2 ($C1^{A}$), 82.0, 81.9, 81.9 (2C), 81.7, 81.6, 78.9, 78.7, 78.4, 77.4 ($C2^{C}$), 76 ($C3^{C}$), 75.5 ($C3^{B}$), 74.7 ($C4^{C}$), 73.6, 73.3, 73.1 (2C), ($C3^{A}$, $C4^{A}$, $C2^{B}$, $C4^{B}$), 72.4 ($C3^{D}$), 71.1, 70.3 ($C2^{A}$, $C2^{D}$), 69.8 ($C4^{D}$), 66.1 ($C5^{A}$), 64.4, 64.1, 64.1 ($C5^{E}$, $C5^{F}$, $C5^{G}$), 63.1, 62.8 ($C5^{C}$, $C5^{D}$), 62.0 ($C5^{B}$). **HRMS:** calcd. for $C_{21}H_{26}O_7SNa^+$ *m/z* 2617.6973; found *m/z* 2617.6916.

Upon counting the carbon signals there are missing one peak from the molecule. Crosspeaks in HSQC, HMBC, and H2BC around δ 76 ppm indicate that this missing peak is hidden by the $CDCl_3$ signal.

It has not been possible to assign every peak in the NMR spectra. However, the tetra-xylose backbone has been assigned, as well as C1, C5, and H1 for the arabinoses, leaving 15 protons and 9 carbons unassigned corresponding to the unassigned signals that the 3 arabinose units would give rise to. Furthermore, the coupling between the anomeric position in the arabinoses and the corresponding carbon or proton in the xylose tetra-backbone can be seen in HMBC. The spectra contains traces of *n*-heptane.

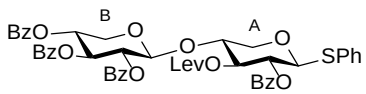
Characterisation of hexasaccharide **phenyl 2,3,4-tri-*O*-benzoyl-β-D-xylopyranosyl-(1→4)-[(1→2)-2,3,5-tri-*O*-benzoyl-α-L-arabinofuranosyl]-[(1→3)-2,3,5-tri-*O*-benzoyl-α-L-arabinofuranosyl]-β-D-xylopyranosyl-(1→4)-2-*O*-benzoyl-β-D-xylopyranosyl-(1→4)-2,3-di-*O*-benzoyl-1-thio-β-D-xylopyranoside**



$[\alpha]_D^{298\,K}$ = -36.1 (*c* 1.0, $CHCl_3$). **^{1}H NMR** (400 MHz, $CDCl_3$) δ 8.17 (d, *J* = 7.2 Hz, 2H, Ar*H*), 8.06 (dd, *J* = 12.2, 7.3 Hz, 4H, Ar*H*), 7.97 – 7.90 (m, 12H, Ar*H*), 7.85 – 7.90 (m, 6H, Ar*H*), 7.68 (d, *J* =

7.3 Hz, 2H, Ar*H*), 7.58 – 7.11 (m, 39H, Ar*H*), 5.79 (s, 1H, H1^{F}), 5.72 (t, *J* = 9.4 Hz, 1H, H3^{D}), 5.65 (s, 1H, H1^{E}), 5.58 (s, 1H, H2^{F}), 5.53 – 5.51 (m, 2H, H2^{E}/H3^{E}, H3^{F}), 5.47 – 5.43 (m, 2H, H3^{A}, H2^{E}/H3^{E}), 5.30 (dd, $J_{2D,3D}$ = 9.4 Hz, $J_{2D,1D}$ = 7.9 Hz, 1H, H2^{D}), 5.27 – 5.19 (m, 2H, H2^{A}, H4^{F}), 5.12 (dt, *J* = 9.6 Hz, $J_{4D,5Deq}$ = 5.6 Hz, 1H, H4^{D}), 5.04 – 4.95 (m, 2H, H5^{F}, H5$^{F'}$), 4.92 – 4.87 (m, 2H, H1^{A}, H2^{B}), 4.70 (dd, $J_{5E,5E'}$ = 11.4 Hz, $J_{5E,4E}$ = 3.5 Hz, 1H, H5^{E}), 4.55 – 4.48 (m, 2H, H1^{D}, H4^{E}), 4.16 (dd, $J_{5E',5E}$ = 12.0 Hz, $J_{5E',4E}$ = 4.5 Hz, 1H, H5$^{E'}$), 4.10 (d, $J_{1B,2B}$ = 7.4 Hz, 1H, H1^{B}), 4.29 (dd, $J_{5Aeq,5Aax}$ = 12.0 Hz, $J_{5Aeq,4A}$ = 4.5 Hz, 1H, H5$^{A}_{eq}$), 4.24 (d, $J_{1C,2C}$ = 7.4 Hz, 1H, H1^{C}), 3.98 – 3.91 (m, 2H, H3^{C}, H5$^{C}_{eq}$), 3.86 – 3.76 (m, 2H, H2^{C}, H4^{C}), 3.72 – 3.67 (m, 2H, H4^{A}, H3^{B}), 3.54 (dd, $J_{5Deq,5Dax}$ = 11.6 Hz, $J_{5Deq,4D}$ = 5.5 Hz, 1H, H5$^{D}_{eq}$), 3.49 (dd, $J_{5Beq,5Bax}$ = 12.0 Hz, $J_{5Beq,4B}$ = 4.7 Hz, 1H, H5$^{B}_{eq}$), 3.41 (dd, $J_{5Aax,5Aeq}$ = 11.9 Hz, $H_{5Aax,4A}$ = 8.7 Hz, 1H, H5$^{A}_{ax}$), 3.34 – 3.30 (m, 2H, H4^{B}, OH), 3.23 – 3.18 (m, 1H, H5$^{D}_{ax}$), 3.10 (dd, $J_{5Bax,5Beq}$ = 11.9 Hz, $J_{5Bax,4B}$ = 8.3 Hz, 1H, H5$^{B}_{ax}$), 2.98 (dd, $J_{5Cax,5Ceq}$ = 11.5 Hz, $J_{5Cax,4C}$ = 9.7 Hz, 1H, H5$^{C}_{ax}$). **APT NMR** (101 MHz, $CDCl_3$) δ 166.5 (1C, C_q), 166.3 (1C, C_q), 166.0 (1C, C_q), 165.8 (1C, C_q), 165.7 (1C, C_q), 165.6 (1C, C_q), 165.5 (1C, C_q), 165.4 (2C, C_q), 165.3 (1C, C_q), 164.9 (1C, C_q), 164.7 (1C, C_q) (*C*O, Bz), 133.8 (1C, Ar*C*H), 133.5 (1C, Ar*C*H), 133.4 (2C, Ar*C*H), 133.3 (2C, Ar*C*H), 133.2 (1C, Ar*C*H), 133.2 (1C, Ar*C*H), 133.1 (1C, Ar*C*H), 132.9 (3C, Ar*C*H), 132.9 (1C, Ar*C*H), 132.5 (1C, Ar*C*), 130.2 (2C, Ar*C*H), 130.1 (1C, Ar*C*), 130.0 (2C, Ar*C*H), 130.0 (4C, Ar*C*H), 129.9 (2C, Ar*C*H), 129.9 (2C, Ar*C*H), 129.8 (3C, Ar*C*H), 129.8 (2C, Ar*C*H), 129.8 (3C, Ar*C*H), 129.7 (1C, Ar*C*), 129.6 (1C, Ar*C*), 129.4 (1C, Ar*C*), 129.4 (1C, Ar*C*), 129.2 (1C, Ar*C*), 129.1 (1C, Ar*C*), 129.1 (2C, Ar*C*H), 129.0 (1C, Ar*C*), 128.9 (1C, Ar*C*), 128.9 (1C, Ar*C*), 128.7 (2C, Ar*C*H), 128.7 (2C, Ar*C*H), 128.5 (4C, Ar*C*H), 128.5 (4C, Ar*C*H), 128.4 (2C, Ar*C*H), 128.3 (2C, Ar*C*H), 128.3 (2C, Ar*C*H), 128.3 (2C, Ar*C*H), 106.7 (C1^{E}), 105.8 (C1^{F}), 102.1 (C1^{C}), 99.6 (2C, C1^{B}, C1^{D}), 86.4 (C1^{A}), 82.1, 82.0 (C4^{E}, C2^{F}), 81.8 (C3^{F}), 81.4 (C4^{F}), 78.7 (C2^{C}), 78.5 (C2^{E}/C3^{E}), 78.1 (C3^{A}), 77.7 (C4^{B}), 76.7 (C3^{C}), 74.1 (C4^{C}), 73.8 (C4^{A}), 72.9 (C2^{E}/C3^{E}), 72.8 (C2^{B}), 72.3 (C3^{D}), 71.8, 71.5 (C3^{B}, C2^{D}), 70.2 (C2^{A}), 69.7 (C4^{D}), 65.6 (C5^{A}), 64.3 (2C, C5^{E}, C5^{F}), 63.0, 62.9 (C5^{C}/C5^{D}), 61.8 (C5^{B}). The spectra contain toluene, *n*-hexane, and H_2O. **HRMS:** calcd. for $C_{120}H_{102}O_{36}SNa^+$ *m/z* 2173.5764; found *m/z* 2173.5717.

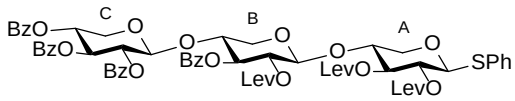
Phenyl 2,3,4-tri-*O*-benzoyl-β-D-xylopyranosyl-(1→4)-2-*O*-benzoyl-3-*O*-levulinoyl-1-thio-β-D-xylopyranoside (110)



General procedure A: Crushed molecular sieves (2.566 g), donor **80** (1.002 g, 1.81 mmol, 1.0 equiv.), AgOTf (0.926 g, 3.60 mmol, 2.0 equiv.), and 4-nitrobenzenesulfenyl chloride (0.342 g, 1.80 mmol, 1.0 equiv.). Activation time

15 min. Acceptor **84** (0.720 g, 1.62 mmol, 0.9 equiv.). Reaction time 50 min. Et_3N (0.40 mL, 5.40 mmol, 3.0 equiv.). Eluent for flash column chromatography (hexane/EtOAc 2.5:1). The product was obtained as white crystals (1.177 g, 82%). $\mathbf{R_f}$(hexane/EtOAc 1:1) 0.42. $[\boldsymbol{\alpha}]_{\boldsymbol{D}}^{\boldsymbol{298\,K}}$ = -60.4 (*c* 1.0, $CHCl_3$). **^{1}H NMR** (400 MHz, $CDCl_3$) δ 8.05 – 8.00 (m, 4H, Ar*H*), 7.97 – 7.92 (m, 4H, Ar*H*), 7.62 – 7.45 (m, 4H, Ar*H*), 7.42 – 7.32 (m, 10H, Ar*H*), 7.28 – 7.25 (m, 3H, Ar*H*), 5.70 (t, J = 6.5 Hz, 1H, $H3^B$), 5.40 (t, $J_{3A,2A}$ = 8.3 Hz, 1H, $H3^A$), 5.28 – 5.23 (m, 2H, $H2^B$, $H4^B$), 5.16 (t, J = 8.4 Hz, 1H, $H2^A$), 4.94 (d, $J_{1B,2B}$ = 4.7 Hz, 1H, $H1^B$), 4.89 (d, $J_{1A,2A}$ = 8.6 Hz, 1H, $H1^A$), 4.48 (dd, $J_{5Beq,5Bax}$ = 12.4, Hz, $J_{5Beq,4B}$ = 3.8 Hz, 1H, $H5^B{}_{eq}$), 4.19 (dd, $J_{5Aeq,5Aax}$ = 12.0, $J_{5Aeq,4A}$ = 4.9 Hz, 1H, $H5^A{}_{eq}$), 3.99 (td, J = 8.7, Hz, $J_{4A,5Aeq}$ = 5.0 Hz, 1H, $H4^A$), 3.75 (dd, $J_{5Bax,5Beq}$ = 12.4 Hz, $J_{5Bax,4B}$ = 6.1 Hz, 1H, $H5^B{}_{ax}$), 3.44 (dd, $J_{5Aax,5Aeq}$ = 12.0 Hz, $J_{5Aax,4A}$ = 9.1 Hz, 1H, $H5^A{}_{ax}$), 2.65 – 2.41 (m, 4H, $COCH_2CH_2$, CH_2CH_2COO), 1.97 (s, 3H, CH_3CO). **^{13}C NMR** (101 MHz, $CDCl_3$) δ 205.9 (1C, C_q, *C*O, Lev), 171.8 (1C, C_q, *C*OO, Lev), 165.6 (1C, C_q), 165.4 (1C, C_q), 165.4 (1C, C_q), 165.2 (1C, C_q) (*C*O, Bz), 133.6 (1C, Ar*C*H), 133.5 (2C, Ar*C*H), 133.5 (1C, Ar*C*H), 132.8 (2C, Ar*C*H), 132.5 (1C, Ar*C*), 130.1 (2C, Ar*C*H), 130.1 (2C, Ar*C*H), 130.1 (2C, Ar*C*H), 130.0 (2C, Ar*C*H), 129.5 (1C, Ar*C*), 129.4 (1C, Ar*C*), 129.2 (1C, Ar*C*), 129.1 (1C, Ar*C*), 129.1 (2C, Ar*C*H), 128.6 (2C, Ar*C*H), 128.6 (2C, Ar*C*H), 128.6 (3C, Ar*C*H), 128.2 (1C, Ar*C*H), 99.4 ($C1^B$), 86.7 ($C1^A$), 74.6 ($C4^A$), 73.2 ($C3^A$), 70.7 ($C2^A$), 70.3 ($C2^B$), 69.6 ($C3^B$), 68.7 ($C4^B$), 66.1 ($C5^A$), 61.1 ($C5^B$), 37.9 ($COCH_2CH_2$), 29.6 (CH_3CO), 28.1 (CH_2CH_2COO). **HRMS:** calcd. for $C_{49}H_{44}O_{14}SNa^+$ *m/z* 911.2344; found *m/z* 911.2340.

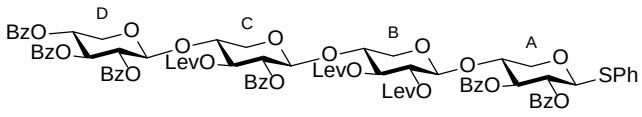
Phenyl 2,3,4-tri-*O*-benzoyl-β-D-xylopyranosyl-(1→4)- 2-*O*-benzoyl-3-*O*-levulinoyl-β-D-xylopyranosyl-(1→4)- 2,3-di-*O*-levulinoyl-1-thio-β-D-xylopyranoside (111):



General procedure A: Crushed molecular sieves (2.560 g), donor **110** (1.001 g, 1.12 mmol, 1.0 equiv.), AgOTf (0.579 g, 2.25 mmol, 2.0 equiv.), and 4-nitrobenzenesulfenyl chloride (0.214 g, 1.12 mmol, 1.0 equiv.). Activation time 40 min. Acceptor **92** (0.447 g, 1.02 mmol, 0.9 equiv.). Reaction time 50 min. Et_3N (0.47 mL, 3.37 mmol, 3.0 equiv.) Eluent for flash column chromatography (hexane/toluene/EtOAc 3:1:4). The product was obtained as white crystals (0.883 g, 71%), furthermore was the donor **110** reisolated (0.119 g, 12%). $\mathbf{R_f}$(hexane/acetone 3:2) 0.48. $[\boldsymbol{\alpha}]_{\boldsymbol{D}}^{\boldsymbol{298\,K}}$ = -60.8 (*c* 1.0, $CHCl_3$). **^{1}H NMR** (400 MHz, $CDCl_3$) δ 8.02 – 7.92 (m, 8H, Ar*H*), 7.61 – 7.43 (m, 6H, Ar*H*), 7.41 – 7.32 (m, 8H, Ar*H*), 7.28 – 7.26 (m, 3H, Ar*H*), 5.70 (t, J = 6.5 Hz, 1H, $H3^C$), 5.31 (t, J = 8.1 Hz, 1H, $H3^B$), 5.28 – 5.24 (m, 2H, $H2^C$, $H4^C$), 5.11 (t, $J_{3A,2A}$ = 8.7 Hz, 1H, $H3^A$), 5.00 (dd, $J_{2B,3B}$ = 8.3 Hz, $J_{2B,1B}$ = 6.4 Hz, 1H, $H2^B$),

4.97 (d, $J_{1C,2C}$ = 4.7 Hz, 1H, H1^{C}), 4.84 (t, $J_{2A,3A}$ = 9.0 Hz, 1H, H2^{A}), 4.61 – 4.58 (m, 2H, H1^{A}, H1^{B}), 4.46 (dd, $J_{5Ceq,5Cax}$ = 12.4 Hz, $J_{5Ceq,4C}$ = 3.8 Hz, 1H, H5$^{C}_{eq}$), 4.06 (dd, $J_{5Beq,5Bax}$ = 11.9 Hz, $J_{5Beq,4B}$ = 4.8 Hz, 1H, H5$^{B}_{eq}$), 4.02 – 3.98 (m, 1H, H4^{B}), 3.94 (dd, $J_{5Aeq,5Aax}$ = 11.9 Hz, $J_{5Aeq,4A}$ = 5.2 Hz, 1H, H5$^{A}_{eq}$), 3.77 – 3.71 (m, 2H, H4^{A}, H5$^{c}_{ax}$), 3.38 (dd, $J_{5Bax,5Beq}$ = 11.9 Hz, $J_{5Bax,4B}$ = 8.1 Hz, 1H, H5$^{B}_{ax}$), 3.21 (dd, $J_{5Aax,5Aeq}$ = 11.8 Hz, $J_{5Aax,4A}$ = 9.8 Hz, 1H, H5$^{A}_{ax}$), 2.85 – 2.39 (m, 12H, COCH_2CH$_2$, CH$_2$CH_2COO), 2.17 (s, 3H), 2.16 (s, 3H), 1.98 (s, 3H) (CH_3CO). **^{13}C NMR** (101 MHz, $CDCl_3$) δ 206.5 (1C, C_q), 206.4 (1C, C_q), 206.0 (1C, C_q) (*C*O, Lev), 171.9 (1C, C_q), 171.7 (1C, C_q), 171.5 (1C, C_q) (*C*OO, Lev), 165.6 (1C, C_q), 165.4 (1C, C_q), 165.2 (1C, C_q), 165.1 (1C, C_q) (*C*O, Bz), 133.5 (1C, Ar*C*H), 133.5 (1C, Ar*C*H), 133.5 (1C, Ar*C*H), 133.5 (1C, Ar*C*H), 132.8 (2C, Ar*C*H), 132.3 (1C, Ar*C*), 130.1 (2C, Ar*C*H), 130.0 (2C, Ar*C*H), 130.0 (2C, Ar*C*H), 129.9 (2C, Ar*C*H), 129.4 (1C, Ar*C*), 129.4 (1C, Ar*C*), 129.3 (1C, Ar*C*), 129.2 (1C, Ar*C*), 129.1 (2C, Ar*C*H), 128.7 (2C, Ar*C*H), 128.5 (4C, Ar*C*H), 128.2 (1C, Ar*C*H), 100.6 (C1^{B}), 99.3 (C1^{C}), 86.5 (C1^{A}), 75.2 (C4^{A}), 74.3 (C4^{B}), 73.4 (C3^{A}), 71.9 (C3^{B}), 71.4 (C2^{B}), 70.2, 70.0 (C2^{A}, C2^{C}), 69.7 (C3^{C}), 68.8 (C4^{C}), 66.6 (C5^{A}), 62.5 (C5^{B}), 61.1 (C5^{C}), 37.9, 37.9, 37.8 (CO*C*H$_2$CH$_2$), 30.0, 29.9, 29.7 (*C*H$_3$CO), 28.1, 28.0, 28.0 (CH$_2$*C*H$_2$COO). **HRMS:** calcd. for $C_{21}H_{26}O_7SNa^+$ *m/z* 1239.3502; found *m/z* 1239.3489.

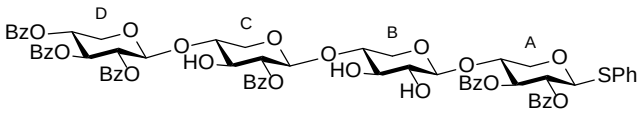
Phenyl 2,3,4-tri-*O*-benzoyl-β-D-xylopyranosyl-(1→4)-2-*O*-benzoyl-3-*O*-levulinoyl-β-D-xylopyranosyl-(1→4)-2,3-di-*O*-levulinoyl-β-D-xylopyranosyl-(1→4)-2,3-di-*O*-benzoyl-1-thio-β-D-xylopyranoside (112)



General procedure A: Crushed molecular sieves (2.147 g), donor **111** (1.207 g, 0.99 mmol, 1.0 equiv.), AgOTf (0.510 g, 1.98 mmol, 2.0 equiv.), 4-nitrobenzenesulfenyl chloride (0.187 g, 0.99 mmol, 1.0 equiv.), and acceptor **81** (0.402 g, 0.89 mmol, 0.9 equiv.). Reaction time 40 min. Et_3N (0.40 mL, 2.97 mmol, 3.0 equiv.) Eluent for flash column chromatography (heptane/toluene/acetone 3:2:2). The product was obtained as white crystals (0.851 g, 61%). $\mathbf{R_f}$(hexane/acetone 3:2) 0.40. $[\boldsymbol{\alpha}]_{\boldsymbol{D}}^{\boldsymbol{298\,K}}$ = -38.2 (*c* 1.0, $CHCl_3$). **^{1}H NMR** (400 MHz, $CDCl_3$) δ 8.01 – 7.91 (m, 12H, Ar*H*), 7.61 (t, *J* = 7.4 Hz, 1H Ar*H*), 7.56 – 7.45 (m, 9H, Ar*H*), 7.41 – 7.32 (m, 10H, Ar*H*), 7.29 – 7.24 (m, 2H, Ar*H*), 7.18 – 7.16 (m, 1H, Ar*H*), 5.70 (t, *J* = 6.5 Hz, 1H, H3^{D}), 5.53 (t, *J* = 7.6 Hz, 1H, H3^{A}), 5.30 – 5.24 (m, 4H, H2^{A}, H3^{C}, H2^{D}, H4^{D}), 5.03 (d, $J_{1A,2A}$ = 7.7 Hz, 1H, H1^{A}), 4.99 – 4.91 (m, 3H, H3^{B}, H2^{C}, H1^{D}), 4.70 (dd, $J_{2B,3B}$ = 8.9 Hz, $J_{2B,1B}$ = 7.0 Hz, 1H, H2^{B}), 4.45 (dd, $J_{5Deq,5Dax}$ = 12.4 Hz, $J_{5Deq,4D}$ = 3.8 Hz, 1H, H5$^{D}_{eq}$), 4.41 (d, $J_{1B,2B}$ = 6.9 Hz, 1H, H1^{B}), 4.36 (d, $J_{1C,2C}$ = 6.5 Hz, 1H, H1^{C}), 4.30 (dd, $J_{5Aeq,5Aax}$ = 12.1 Hz, $J_{5Aeq,4A}$ =

4.6 Hz, 1H, H5$^{A}_{eq}$), 4.01 – 3.86 (m, 3H, H4^{A}, H4^{C}, H5^{C}), 3.74 (dd, $J_{5Dax,5Deq}$ = 12.4 Hz, $J_{5Dax,4D}$ = 6.1 Hz, 1H, H5$^{D}_{ax}$), 3.58 (dd, $J_{5Aax,5Aeq}$ = 12.1 Hz, $J_{5Aax,4A}$ = 8.2 Hz, 1H, H5$^{A}_{ax}$), 3.49 – 3.44 (m, 1H, H4^{B}), 3.34 – 3.28 (m, 2H, H5$^{B}_{eq}$, H5^{C}), 2.89 (dd, $J_{5Bax,5Beq}$ = 12.0 Hz, $J_{5Bax,4B}$ = 9.4 Hz, 1H, H5$^{B}_{ax}$), 2.71 – 2.35 (m, 12H, COCH_2CH$_2$, CH$_2$CH_2COO), 2.14 (d, 6H), 1.97 (s, 3H) (CH_3CO). **^{13}C NMR** (101 MHz, $CDCl_3$) δ 206.4 (1C, C_q), 206.4 (1C, C_q), 205.9 (1C, C_q) (*C*O, Lev), 171.8 (1C, C_q), 171.7 (1C, C_q), 171.3 (1C, C_q) (*C*OO, Lev), 165.6 (1C, C_q), 165.4 (1C, C_q), 165.3 (2C, C_q), 165.2 (1C, C_q), 165.0 (1C, C_q) (*C*O, Bz), 133.5 (1C, Ar*C*H), 133.5 (1C, Ar*C*H), 133.5 (1C, Ar*C*H), 133.5 (1C, Ar*C*H), 133.3 (1C, Ar*C*H), 133.2 (1C, Ar*C*H), 133.0 (1C, Ar*C*), 132.6 (2C, Ar*C*H), 130.1 (3C, Ar*C*H), 130.0 (2C, Ar*C*H), 130.0 (2C, Ar*C*H), 129.9 (2C, Ar*C*H), 129.8 (2C, Ar*C*H), 129.6 (1C, Ar*C*), 129.5 (1C, Ar*C*), 129.4 (1C, Ar*C*), 129.4 (1C, Ar*C*), 129.2 (1C, Ar*C*), 129.2 (1C, Ar*C*), 129.1 (1C, Ar*C*), 129.1 (2C, Ar*C*H), 128.6 (2C, Ar*C*H), 128.6 (3C, Ar*C*H), 128.5 (4C, Ar*C*H), 128.5 (2C, Ar*C*H), 128.4 (2C, Ar*C*H), 128.3 (1C, Ar*C*H), 128.1 (1C, Ar*C*H), 100.9 (C1^{B}), 100.3 (C1^{C}), 99.2 (C1^{D}), 86.6 (C1^{A}), 75.4 (C4^{A}), 74.8 (C4^{B}), 74.4 (C4^{C}), 72.7 (C3^{A}), 72.1, 71.9 (C3^{B}, C2^{A}/C3^{C}/C2^{D}), 71.3, 71.2 (C2^{B}, C2^{C}), 70.4, 70.2 (C2^{A}/C3^{C}/C2^{D}), 69.6 (C3^{D}), 68.7 (C4^{D}), 65.5 (C5^{A}), 62.6, 62.5 (C5^{B}, C5^{C}), 61.1 (C5^{D}), 37.8, 37.8, 37.7 (CO*C*H$_2$CH$_2$), 30.0, 29.9, 29.6 (*C*H$_3$CO), 28.0, 27.9, 27.9 (CH$_2$*C*H$_2$COO). **HRMS:** calcd. for $C_{83}H_{80}O_{28}SNa^+$ *m/z* 1579.4449; found *m/z* 1579.4430.

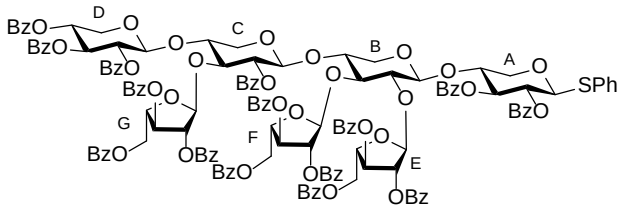
Phenyl 2,3,4-tri-*O*-benzoyl-β-D-xylopyranosyl-(1→4)-2-*O*-benzoyl-β-D-xylopyranosyl-(1→4)-β-D-xylopyranosyl-(1→4)-2,3-di-*O*-benzoyl-1-thio-β-D-xylopyranoside (113)



General procedure B: Tetrasaccharide **112** (0.507 g, 0.33 mmol, 1.0 equiv.), and 50% solution of hydrazine hydrate (0.20 mL, 3.21 mmol, 10.0 equiv.). Reaction time 20 min. Eluent for TLC (hexane/toluene/acetone 2:1:2). Acetone (7.0 mL, 96.3 mmol, 300 equiv.). The product was a white powder (0.400 g, 97%). **R$_f$**(hexane/acetone 3:2) 0.24. $[\boldsymbol{\alpha}]_{\boldsymbol{D}}^{\boldsymbol{298\ K}}$ = -15.4 (*c* 1.0, $CHCl_3$). **^{1}H NMR** (400 MHz, $CDCl_3$) δ 8.04 – 7.92 (m, 12H, Ar*H*), 7.63 (t, *J* = 7.4 Hz, 1H, Ar*H*), 7.57 – 7.26 (m, 22H, Ar*H*), 5.84 (t, $J_{3D,2D}$ = 8.8 Hz, 1H, H3^{D}), 5.57 (t, $J_{3A,2A}$ = 7.5 Hz, 1H, H3^{A}), 5.43 (dd, $J_{2D,3D}$ = 8.7 Hz, $J_{2D,1D}$ = 7.1 Hz, 1H, H2^{D}), 5.38 – 5.30 (m, 2H, H2^{A}, H4^{D}), 5.10 (d, $J_{1A,2A}$ = 7.4 Hz, 1H, H1^{A}), 5.05 (t, *J* = 8.2 Hz, 1H, H2^{C}), 4.88 (d, $J_{1D,2D}$ = 6.9 Hz, 1H, H1^{D}), 4.42 – 4.36 (m, 2H, H5^{A}, H5^{D}), 4.33 (d, *J* = 7.6 Hz, 1H, H1^{C}), 4.27 (d, *J* = 6.7 Hz, 1H, H1^{B}), 3.97 – 3.92 (m, 1H, H4^{A}), 3.89 – 3.81 (m, 3H, H3^{C}, H4^{C}, H5^{C}), 3.73 – 3.62 (m, 4H, H5^{A}', H5^{D}', OH), 3.45 (t, *J* = 8.1 Hz, 1H, H3^{B}), 3.33 – 3.19 (m, 4H, H2^{B}, H4^{B}, H5$^{B}_{eq}$, H5^{C}'), 3.05 (bs, 1H, OH) 2.89 (dd, $J_{5Bax,5Beq}$ = 11.5 Hz, $J_{5bax,4B}$ = 9.2 Hz, 1H, H5$^{B}_{ax}$). **^{13}C NMR** (101 MHz, $CDCl_3$) δ

165.6 (1C, C_q), 165.6 (1C, C_q), 165.6 (1C, C_q), 165.6 (1C, C_q), 165.3 (1C, C_q), 165.2 (1C, C_q) (*C*O, Bz), 133.8 (1C, Ar*C*H), 133.7 (1C, Ar*C*H), 133.5 (2C, Ar*C*H), 133.4 (1C, Ar*C*H), 133.3 (1C, Ar*C*H), 133.0 (1C, Ar*C*), 132.6 (3C, Ar*C*H), 130.0 (4C, Ar*C*H), 129.9 (4C, Ar*C*H), 129.9 (2C, Ar*C*H), 129.8 (2C, Ar*C*H), 129.6 (1C, Ar*C*), 129.5 (1C, Ar*C*), 129.5 (1C, Ar*C*), 129.1 (2C, Ar*C*H), 129.0 (1C, Ar*C*), 129.0 (1C, Ar*C*), 128.9 (1C, Ar*C*), 128.7 (2C, Ar*C*H), 128.7 (2C, Ar*C*H), 128.6 (2C, Ar*C*H), 128.5 (2C, Ar*C*H), 128.5 (2C, Ar*C*H), 128.4 (2C, Ar*C*H), 128.2 (1C, Ar*C*H), 102.1 ($C1^B$, $J_{C\text{-}H}$ = 160 Hz), 101.8 ($C1^D$, $J_{C\text{-}H}$ = 163 Hz), 101.4 ($C1^C$, $J_{C\text{-}H}$ = 163 Hz), 86.6 ($C1^A$, $J_{C\text{-}H}$ = 159 Hz), 80.2 ($C4^C$), 79.0 ($C4^B$), 74.1 ($C4^A$), 73.3, 73.1, 73.0 ($C3^B$, $C2^C$, $C3^C$), 72.5, 72.3 ($C3^A$, $C2^B$), 71.4, 71.2 ($C2^D$, $C3^D$), 70.2 ($C2^A$), 69.4 ($C4^D$), 65.1 ($C5^a$), 63.4 ($C5^C$), 62.8 ($C5^D$), 62.0 ($C5^B$). **HRMS:** calcd. for $C_{68}H_{62}O_{22}SNa^+$ *m/z* 1285.3345; found *m/z* 1285.3334.

Phenyl 2,3,4-tri-*O*-benzoyl-β-D-xylopyranosyl-(1→4)-[(1→3)-2,3,5-tri-*O*-benzoyl-α-L-arabinofuranosyl]- 2-*O*-benzoyl-β-D-xylopyranosyl-(1→4)-[(1→2)-2,3,5-tri-*O*-benzoyl-α-L-arabinofuranosyl]-[(1→3)-2,3,5-tri-*O*-benzoyl-α-L-arabinofuranosyl]-β-D-xylopyranosyl-(1→4)-2,3-di-*O*-benzoyl-1-thio-β-D-xylopyranoside (114)

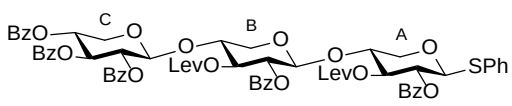


General procedure C: Acceptor **113** (0.201 g, 0.16 mmol, 1.0 equiv.), donor **98** (0.401 g, 0.63 mmol, 3.0 equiv.), and TMSOTf in dry CH_2Cl_2 (0.16 mL, c = 0.10 mmol/mL, 0.1 equiv.). Reaction time: 1 h. Et_3N (0.02 mL, 0.11 mmol, 0.7 equiv.). Eluent TLC analysis (hexane/toluene/acetone 2:1:1) and eluent flash column chromatography (hexane/toluene/acetone 3:2:1→2:1:1). The product was white crystals (0.393 g, 95%). $\mathbf{R_f}$(hexane/toluene/acetone 2:1:1) 0.42. $[\boldsymbol{\alpha}]_{\boldsymbol{D}}^{\boldsymbol{298\,K}}$ = -52.5 (*c* 1.0, $CHCl_3$). **^{1}H NMR** (400 MHz, $CDCl_3$) δ 8.18 (d, *J* = 7.3 Hz, 2H, Ar*H*), 8.09 – 7.73 (m, 28H, Ar*H*), 7.65 (d, *J* = 7.3 Hz, 2H, Ar*H*), 7.58 – 7.05 (m, 48H, Ar*H*), 5.76 (s, 1H, $H1^F$), 5.69 (t, *J* = 9.1 Hz, 1H, $H3^D$), 5.65 (s, 1H, $H1^E$), 5.54 – 5.47 (m, 6H, $H3^A$), 5.34 – 5.34 (m, 1H, $H2^E$), 5.25 (s, 1H, $H1^G$), 5.25 – 5.20 (m, 2H, $H2^A$, $H2^D$), 5.19 – 5.13 (m, 1H, $H4^D$), 5.10 – 5.06 (m, 2H), 5.04 – 4.96 (m, 3H, $H2^C$), 4.94 – 4.85 (m, 3H, $H1^A$), 4.79 (dd, *J* = 11.4 Hz, *J* = 3.0 Hz, 1H), 4.71 – 4.62 (m, 2H), 4.31 (d, $J_{1B,2B}$ = 6.9 Hz, 1H, $H1^B$), 4.17 – 4.13 (m, 2H, $H5^A$, $H1^D$), 3.88 – 3.81 (m, 4H, $H3^B$, $H1^C$, $H3^C$, $H5^D$), 3.79 – 3.74 (m, 1H, $H4^A$), 3.65 (dd, $J_{2B,3B}$ = 9.5 Hz, $J_{2B,1B}$ = 7.0 Hz, 1H, $H2^B$), 3.53 – 3.47 (m, 2H, $H5^{A'}$, $H4^C$), 3.32 – 3.26 (m, 3H, $H4^B$, $H5^B$, $H5^{D'}$), 3.03 (dd, $J_{5C,5C'}$ = 11.8 Hz, $J_{5C,4C}$ = 5.1 Hz, 1H, $H5^C$), 2.73 – 2.67 (m, 1H, $H5^{B'}$), 2.53 (t, *J* = 11.1 Hz, 1H, $H5^{C'}$). **^{13}C NMR** (101 MHz, $CDCl_3$) δ 166.5 (1C, C_q), 166.3 (1C, C_q), 166.3 (1C,

C_q), 166.0 (1C, C_q), 165.8 (1C, C_q), 165.8 (1C, C_q), 165.7 (1C, C_q), 165.6 (1C, C_q), 165.3 (1C, C_q), 165.2 (2C, C_q), 165.0 (1C, C_q), 165.0 (1C, C_q), 164.5 (1C, C_q), 164.2 (1C, C_q) (*C*O, Bz), 133.6 (1C, Ar*C*H), 133.6 (1C, Ar*C*H), 133.6 (1C, Ar*C*H), 133.5 (1C, Ar*C*H), 133.4 (1C, Ar*C*H), 133.4 (1C, Ar*C*H), 133.3 (1C, Ar*C*H), 133.2 (1C, Ar*C*H), 133.1 (1C, Ar*C*H), 133.0 (2C, Ar*C*H), 132.9 (1C, Ar*C*H), 132.8 (1C, Ar*C*H), 132.2 (2C, Ar*C*H), 130.3 (1C, Ar*C*H), 130.1 (2C, Ar*C*H), 130.1 (2C, Ar*C*H), 130.0 (2C, Ar*C*H), 129.9 (3C, Ar*C*H), 129.9 (3C, Ar*C*H), 129.8 (2C, Ar*C*H), 129.8 (4C, Ar*C*H), 129.7 (2C, Ar*C*H), 129.7 (1C, Ar*C*), 129.5 (1C, Ar*C*), 129.2 (2C, Ar*C*), 129.1 (1C, Ar*C*), 129.1 (1C, Ar*C*), 129.0 (1C, Ar*C*), 129.0 (1C, Ar*C*), 129.0 (2C, Ar*C*H), 128.9 (1C, Ar*C*), 128.9 (1C, Ar*C*), 128.6 (2C, Ar*C*H), 128.6 (3C, Ar*C*H), 128.4 (2C, Ar*C*H), 128.4 (2C, Ar*C*H), 128.3 (3C, Ar*C*H), 128.3 (2C, Ar*C*H), 128.2 (3C, Ar*C*H), 127.9 (1C, Ar*C*H), 106.9 ($C1^E$), 106.2 ($C1^G$), 105.4 ($C1^F$), 101.7 ($C1^B$), 99.8 ($C1^D$), 99.5 ($C1^C$), 86.2 ($C1^A$), 82.6 ($C2^F$), 82.3, 82.2, 81.8, 81.8, 80.7, 79.7 ($C2^B$), 78.6, 78.1, 77.8, 76.0 ($C3^B$/$C3^C$), 75.8 ($C3^B$/$C3^C$), 74.6 ($C4^C$), 73.8 ($C4^B$), 73.6 ($C4^A$), 73.1 ($C2^C$), 71.9 ($C3^D$), 71.8 ($C3^A$), 71.1 ($C2^D$), 70.2 ($C2^A$), 69.7 ($C4^D$), 64.6 ($C5^A$), 64.2, 64.0, 63.9 ($C5^E$, $C5^F$, $C5^G$), 63.0, 62.8 ($C5^C$, $C5^D$), 62.3 ($C5^B$). **HRMS:** calcd. for $C_{146}H_{122}O_{43}S(Na^+)_2$ *m/z* 2640.6865; found *m/z* 2640.6926.

It has not been possible to assign every peak in the NMR spectra. However, the tetra-xylose backbone has been assigned, as well as C1, C5, and H1 for the arabinoses, leaving 15 protons and 9 carbons unassigned corresponding to the unassigned signals that the 3 arabinose units would give rise to. Furthermore, the coupling between the anomeric position in the arabinoses and the corresponding carbon or proton in the xylose tetra-backbone can be seen in HMBC. The spectra contain traces of toluene and *n*-heptane.

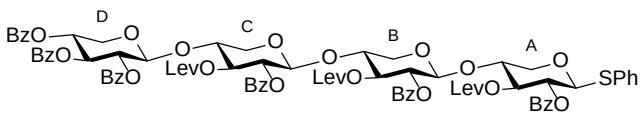
Phenyl 2,3,4-tri-*O*-benzoyl-β-D-xylopyranosyl-(1→4)-2-*O*-benzoyl-3-*O*-levulinoyl-β-D-xylopyranosyl-(1→4)-2-*O*-benzoyl-3-*O*-levulinoyl-1-thio-β-D-xylopyranoside (115):



General procedure A: Crushed molecular sieves (1.489 g), donor **110** (0.499 g, 0.56 mmol, 1.0 equiv.), AgOTf (0.289 g, 1.12 mmol, 2.0 equiv.), and 4-nitrobenzenesulfenyl chloride (0.106 g, 0.56 mmol, 1.0 equiv.). Activation time: 35 min. Acceptor **84** (0.226 g, 0.51 mmol, 0.9 equiv.). Reaction time 45 min. Et_3N (0.23 mL, 1.68 mmol, 3.0 equiv.) Eluent for flash column chromatography (heptane/EtOAc 4:3). The product was obtained as white crystals (0.446 g, 72%). $\mathbf{R_f}$(heptane/EtOAc 1:1) 0.37. $[\boldsymbol{\alpha}]_{\boldsymbol{D}}^{\boldsymbol{298\ K}}$ = - 57.7 (*c* 1.0,

$CHCl_3$). **^{1}H NMR** (400 MHz, $CDCl_3$) δ 8.05 – 7.95 (m, 10H, Ar*H*), 7.63 – 7.46 (m, 9H, Ar*H*), 7.43 – 7.35 (m, 8H, Ar*H*), 7.29 – 7.25 (m, 3H, Ar*H*), 5.73 (t, J = 6.6 Hz, 1H, $H3^C$), 5.37 – 5.26 (m, 4H, $H3^A$, $H3^B$, $H2^C$, $H4^C$), 5.13 (t, $J_{2A,1A}$ = 8.3 Hz, 1H, $H2^A$), 5.05 (dd, $J_{2B,3B}$ = 8.3 Hz, $J_{2B,1B}$ = 6.5 Hz, 1H, $H2^B$), 4.96 (d, $J_{1C,2C}$ = 4.7 Hz, 1H, $H1^C$), 4.86 (d, $J_{1A,2A}$ = 8.4 Hz, 1H, $H1^A$), 4.67 (d, $J_{1B,2B}$ = 6.4 Hz, 1H, $H1^B$), 4.48 (dd, $J_{5Ceq,5Cax}$ = 12.4 Hz, $J_{5Ceq,4C}$ = 3.9 Hz, 1H, $H5^C{}_{eq}$), 4.11 – 4.04 (m, 2H, $H5^A{}_{eq}$, $H5^B{}_{eq}$), 3.98 – 3.97 (m, 1H, $H4^B$), 3.88 – 3.82 (m, 1H, $H4^A$), 3.77 (dd, $J_{5Cax,5Ceq}$= 12.4 Hz, $J_{5Cax,4C}$ = 6.2 Hz, 1H, $H5^C{}_{ax}$), 3.42 (dd, $J_{5Bax,5Beq}$ = 12.1 Hz, $J_{5Bax,4B}$ = 8.4 Hz, 1H, $H5^B{}_{ax}$), 3.36 (dd, $J_{5Aax,5Aeq}$ = 11.9 Hz, $J_{5Aax,4A}$ = 9.0 Hz, 1H, $H5^A{}_{ax}$), 2.66 - 2.39 (m, 8H, $COCH_2CH_2$, CH_2CH_2COO), 2.06 (s, 3H), 2.01 (s, 3H) (CH_3CO). **^{13}C NMR** (101 MHz, $CDCl_3$) δ 206.0 (1C, C_q), 205.9 (1C, C_q) (*C*O, Lev), 171.8 (1C, C_q), 171.6 (1C, C_q) (*C*OO, Lev), 165.6 (1C, C_q), 165.3 (1C, C_q), 165.3 (1C, C_q), 165.1 (2C, C_q) (*C*O, Bz), 133.5 (1C, Ar*C*H), 133.5 (1C, Ar*C*H), 133.4 (1C, Ar*C*H), 133.4 (1C, Ar*C*H), 132.6 (2C, Ar*C*H), 132.6 (1C, Ar*C*), 130.1 (2C, Ar*C*H), 130.0 (2C, Ar*C*H), 130.0 (2C, Ar*C*H), 129.9 (2C, Ar*C*H), 129.9 (2C, Ar*C*H), 129.5 (1C, Ar*C*), 129.4 (1C, Ar*C*), 129.4 (1C, Ar*C*), 129.2 (1C, Ar*C*), 129.1 (1C, Ar*C*), 129.0 (2C, Ar*C*H), 128.6 (2C, Ar*C*H), 128.5 (4C, Ar*C*H), 128.5 (4C, Ar*C*H), 128.1 (1C, Ar*C*H), 100.5 ($C1^B$), 99.2 ($C1^C$), 86.5 ($C1^A$), 74.9 ($C4^A$), 74.3 ($C4^B$), 73.0 ($C3^A$), 71.9 ($C3^B$), 71.4 ($C2^B$), 70.4, 70.2 ($C2^A$, $C2^C$), 69.7 ($C3^C$), 68.7 ($C4^C$), 66.1 ($C5^A$), 62.4 ($C5^B$), 61.1 ($C5^C$), 37.8, 37.8 ($CO*C*H_2CH_2$), 29.8, 29.6 ($*C*H_3CO$), 28.0, 28.0 (CH_2*C*H_2COO). **HRMS:** calcd. for $C_{66}H_{62}O_{21}SNa^+$ *m/z* 1245.3396; found *m/z* 1245.3379.

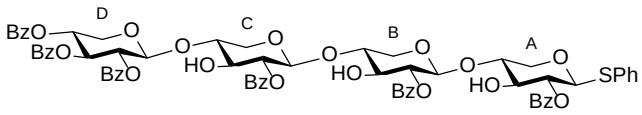
Phenyl 2,3,4-tri-*O*-benzoyl-β-D-xylopyranosyl-(1→4)-2-*O*-benzoyl-3-*O*-levulinoyl-β-D-xylopyranosyl-(1→4)-2-*O*-benzoyl-3-*O*-levulinoyl-β-D-xylopyranosyl-(1→4)-2-*O*-benzoyl-3-*O*-levulinoyl-1-thio-β-D-xylopyranoside (116)



General procedure A: Crushed molecular sieves (1.491 g), donor **115** (0.794 g, 0.65 mmol, 1.0 equiv.), AgOTf (0.332 g, 1.29 mmol, 2.0 equiv.), 4-nitrobenzenesulfenyl chloride (0.123 g, 0.65 mmol, 1.0 equiv.), and acceptor **84** (0.262 g, 0.59 mmol, 0.9 equiv.). Reaction time 60 min. Et_3N (0.27 mL, 1.95 mmol, 3.0 equiv.) Eluent for flash column chromatography (heptane/acetone 2:1). The product was obtained as white crystals (0.670 g, 73%). $\mathbf{R_f}$(hexane/acetone 1:1) 0.34. $[\boldsymbol{\alpha}]_{\boldsymbol{D}}^{\boldsymbol{298\,K}}$ = -46.6 (*c* 1.0, $CHCl_3$). **^{1}H NMR** (400 MHz, $CDCl_3$) δ 8.01 – 7.92 (m, 12H, Ar*H*), 7.60 – 7.31 (m, 20H, Ar*H*), 7.25 – 7.21 (m, 3H, Ar*H*), 5.70 (t, J = 6.5 Hz, 1H, $H3^D$), 5.32 (t, $J_{3B,2B}$ = 8.2 Hz, 1H, $H3_B$), 5.28 – 5.24 (m, 3H, $H3^A$, $H2^D$, $H4^D$), 5.20 (t, $J_{3C,2C}$ = 8.0 Hz, 1H, $H3^C$), 5.08 (t, $J_{2A,1A}$ = 8.3 Hz, 1H, $H2^A$), 5.04 (dd, $J_{2B,3B}$ = 8.4

Hz, $J_{2B,1B}$ = 6.5 Hz, 1H, H2^{B}), 4.97 (dd, $J_{2C,3C}$ = 8.3 Hz, $J_{2C,1C}$ = 6.4 Hz, 1H, H2^{C}), 4.93 (d, $J_{1D,2D}$ = 4.7 Hz, 1H, H1^{D}), 4.81 (d, $J_{1A,2A}$ = 8.4 Hz, 1H, H1^{A}), 4.64 (d, $J_{1B,2B}$ = 6.5 Hz, 1H, H1^{B}), 4.57 (d, $J_{1C,2C}$ = 6.3 Hz, 1H, H1^{C}), 4.46 (dd, $J_{5Deq,5Dax}$ = 12.4 Hz, $J_{5Deq,4D}$ = 3.9 Hz, 1H, H5$^{D}_{eq}$), 4.05 – 4.00 (m, 2H, H5^{A}, H5$^{B}_{eq}$), 3.95 – 3.89 (m, 2H, H4^{B}, H5^{C}), 3.80 – 3.72 (m, 3H, H4^{A}, H4^{C}, H5$^{D}_{ax}$), 3.38 (dd, $J_{5Bax,5Beq}$ = 12.1 Hz, $J_{5Bax,4B}$ = 8.5 Hz, 1H, H5$^{B}_{ax}$), 3.33 – 3.25 (m, 2H, H5^{A}', H5^{C}'), 2.66 – 2.32 (m, 12H, COCH_2CH$_2$, CH$_2$CH_2COO) 2.03 (s, 3H), 2.01 (s, 3H), 1.97 (s, 3H) (CH_3CO). **^{13}C NMR** (101 MHz, $CDCl_3$) δ 206.1 (1C, C_q), 206.0 (1C, C_q), 205.9 (1C, C_q) (*C*O, Lev), 171.8 (1C, C_q), 171.7 (1C, C_q), 171.5 (1C, C_q) (*C*OO, Lev), 165.6 (1C, C_q), 165.3 (1C, C_q), 165.2 (1C, C_q), 165.2 (1C, C_q), 165.1 (1C, C_q), 165.1 (1C, C_q) (*C*O, Bz), 133.5 (1C, Ar*C*H), 133.5 (2C, Ar*C*H), 133.4 (2C, Ar*C*H), 132.6 (2C, Ar*C*H), 132.5 (1C, Ar*C*), 130.1 (2C, Ar*C*H), 130.0 (2C, Ar*C*H), 130.0 (2C, Ar*C*H), 129.9 (3C, Ar*C*H), 129.9 (3C, Ar*C*H), 129.5 (1C, Ar*C*), 129.4 (2C, Ar*C*), 129.4 (1C, Ar*C*), 129.2 (1C, Ar*C*), 129.1 (1C, Ar*C*), 129.0 (2C, Ar*C*H) 128.6 (3C, Ar*C*H), 128.5 (4C, Ar*C*H), 128.5 (4C, Ar*C*H), 128.1 (1C, Ar*C*H), 100.4 (2C, C1^{B}, C1^{C}), 99.2 (C1^{D}), 86.5 (C1^{A}), 74.7 (C4^{A}), 74.5, 74.3 (C4^{B}, C4^{C}), 72.9 (C3^{A}), 71.9 (C3^{B}), 71.6 (C3^{C}), 71.4 (C2^{B}), 71.1 (C2^{C}), 70.4 (C2^{A}), 70.2 (C2^{D}), 69.6 (C3^{D}), 68.7 (C4^{D}), 66.0 (C5^{A}), 62.4 (2C, C5^{B}, C5^{C}), 61.0 (C5^{D}), 37.8, 37.8, 37.8 (CO*C*H$_2$CH$_2$), 29.8, 29.8, 29.6 (*C*H$_3$CO), 28.0, 28.0, 27.9 (CH$_2$*C*H$_2$COO). The spectra contain traces of *n*-heptane. **HRMS:** calcd. for $C_{83}H_{80}O_{28}SNa^+$ *m/z* 1579.4449; found *m/z* 1579.4427.

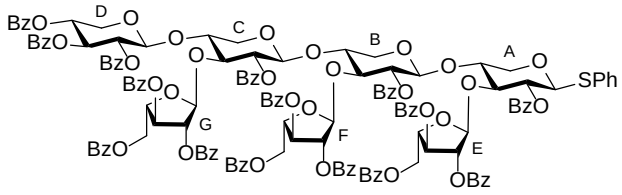
Phenyl 2,3,4-tri-*O*-benzoyl-β-D-xylopyranosyl-(1→4)-2-*O*-benzoyl-β-D-xylopyranosyl-(1→4)-2-*O*-benzoyl-β-D-xylopyranosyl-(1→4)-2-*O*-benzoyl-1-thio-β-D-xylopyranoside (117)



General procedure B: Tetrasaccharide **116** (0.146 g, 0.094 mmol, 1.0 equiv.), and 50% solution of hydrazine hydrate (59 µL, 0.94 mmol, 10.0 equiv.). Reaction time 3 h. Eluent for TLC (heptane/EtOAc 1:1). Acetone (2.1 mL, 28.6 mmol, 304 equiv.). The product obtained was a white powder (0.113 g, 96%). $\mathbf{R_f}$(heptane/acetone 1:1) 0.37. $[\boldsymbol{\alpha}]_{\boldsymbol{D}}^{\boldsymbol{298\,K}}$ = - 36.6 (*c* 1.0, $CHCl_3$). **^{1}H NMR** (400 MHz, $CDCl_3$) δ 8.06 – 8.00 (m, 6H, Ar*H*), 7.96 – 7.90 (m, 6H, Ar*H*), 7.62 – 7.32 (m, 22H, Ar*H*), 7.24 – 7.20 (m, 1H, Ar*H*), 5.82 (t, *J* = 8.9 Hz, 1H, H3^{D}), 5.42 (dd, $J_{2D,3D}$ = 9.0 Hz, $J_{2D,1D}$ = 7.0 Hz, 1H, H1^{D}), 5.37 (td, $J_{4D,3D}$ = 8.9 Hz, *J* = 4.5 Hz, 1H, H4^{D}), 5.11 – 5.07 (m, 1H, H2^{C}), 5.05 – 5.00 (m, 2H, H2^{A}, H2^{B}), 4.83 (d, $J_{1D,2D}$ = 7.0 Hz, 1H, H1^{D}), 4.64 (d, $J_{1A,2A}$ = 9.9 Hz, 1H, H1^{A}), 4.50 (d, J_{1CH2C} = 7.9 Hz, 1H, H1^{C}), 4.47 – 4.43 (m, 2H, H1^{B}, H5^{D}), 3.87 – 3.77 (m, 4H, H3^{A}, H3^{C}, H4^{C}, H5^{C}), 3.75 – 3.55 (m, 7H, H3^{A}, H4^{A},

H3^{B}, H4^{B}, H5^{B}, H5$^{D'}$), 3.31 – 3.15 (m, 3H, H5$^{A'}$, H5$^{B'}$, H5$^{C'}$), 1.65 (bs, 2H, OH). **^{13}C NMR** (101 MHz, $CDCl_3$) δ 165.6 (2C, C_q), 165.5 (1C, C_q), 165.5 (1C, C_q), 165.4 (1C, C_q), 165.2 (1C, C_q) (*C*O, Bz), 133.8 (1C, Ar*C*H), 133.7 (1C, Ar*C*H), 133.6 (1C, Ar*C*H), 133.5 (1C, Ar*C*H), 133.5 (1C, Ar*C*H), 133.3 (1C, Ar*C*H), 132.8 (3C, Ar*C*H), 132.3 (1C, Ar*C*), 130.0 (2C, Ar*C*H), 129.9 (2C, Ar*C*H), 129.9 (2C, Ar*C*H), 129.9 (2C, Ar*C*H), 129.8 (2C, Ar*C*H), 129.7 (4C, Ar*C*H), 129.5 (1C, Ar*C*), 129.4 (1C, Ar*C*), 129.1 (1C, Ar*C*H), 129.0 (1C, Ar*C*H), 129.0 (3C, Ar*C*H), 128.9 (1C, Ar*C*), 128.9 (1C, Ar*C*), 128.9 (1C, Ar*C*), 128.8 (2C, Ar*C*H), 128.7 (2C, Ar*C*H), 128.6 (2C, Ar*C*H), 128.6 (2C, Ar*C*H), 128.5 (2C, Ar*C*H), 128.5 (2C, Ar*C*H), 128.3 (1C, Ar*C*H), 128.1 (1C, Ar*C*H), 102.2, 102.1 (C1^{B}, C1^{C}, $J_{C\text{-}H}$ = 157 Hz, $J_{C\text{-}H}$ = 163 Hz), 101.8 (C1^{D}, $J_{C\text{-}H}$ = 162 Hz), 86.6 (C1^{A}, $J_{C\text{-}H}$ = 155 Hz), 80.6 (C4^{A}), 80.3 (C4^{B}), 80.1 (C4^{C}), 74.8 (C3^{A}), 73.2, 73.1, 73.0, 73.0 (C2^{B}, C3^{B}, C2^{C}, C3^{C}), 72.0 (C2^{A}), 71.3 (C3^{D}), 71.1 (C2^{D}), 69.3 (C4^{D}), 67.1 (C5^{A}), 63.5 (2C, C5^{B}, C5^{C}), 62.8 (C5^{D}). **HRMS:** calcd. for $C_{68}H_{62}O_{22}SNa^+$ *m/z* 1285.3346; found *m/z* 1285.3346.

Phenyl 2,3,4-tri-*O*-benzoyl-β-D-xylopyranosyl-(1→4)-[(1→3)-2,3,5-tri-*O*-benzoyl-α-L-arabinofuranosyl]-2-*O*-benzoyl-β-D-xylopyranosyl-(1→4)-[(1→3)-2,3,5-tri-*O*-benzoyl-α-L-arabinofuranosyl]-2-*O*-benzoyl-β-D-xylopyranosyl-(1→4)-[(1→3)-2,3,5-tri-*O*-benzoyl-α-L-arabinofuranosyl]-2-*O*-benzoyl-1-thio-β-D-xylopyranoside (118)

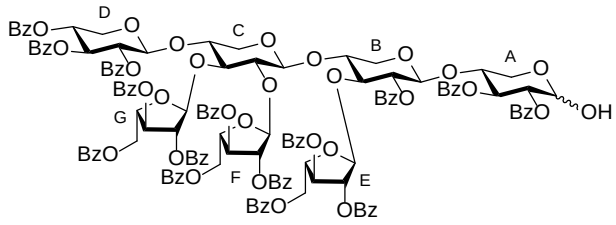


General procedure C: Acceptor **117** (0.054 g, 0.043 mmol, 1.0 equiv.), donor **98** (0.109 g, 0.17 mmol, 4.0 equiv.), and TMSOTf in dry CH_2Cl_2 (0.11 mL, c = 0.04 mmol/mL, 0.1 equiv.). Reaction time: 75 min. Et_3N (0.02 mL, 0.11 mmol, 2.6 equiv.). Eluent TLC (heptane/acetone 1:1) and flash column chromatography (heptane/acetone 2:1). The product **118** was isolated as white crystals (0.100 g, 90%). $\mathbf{R_f}$(heptane/acetone 1:1) 0.41. $[\boldsymbol{\alpha}]_{\boldsymbol{D}}^{\boldsymbol{298\ K}}$ = -46.9 (*c* 1.0, $CHCl_3$). **^{1}H NMR** (400 MHz, $CDCl_3$) δ 8.18 – 8.16 (m, 2H, Ar*H*), 8.09 – 7.87 (m, 28H, Ar*H*), 7.83 (t, *J* = 7.4 Hz, 1H, Ar*H*), 7.64 – 7.16 (m, 49H, Ar*H*), 5.71 (t, $J_{3D,4D}$ = 9.3 Hz, 1H, H3^{D}), 5.52 (d, *J* = 5.5 Hz, 1H), 5.49 (d, *J* = 5.3 Hz, 1H), 5.45 (d, *J* = 4.6 Hz, 1H), 5.43 (s, 1H), 5.37 (d, *J* = 1.0 Hz, 1H), 5.35 – 5.34 (m, 2H), 5.30 – 5.28 (m, 2H, H2^{D}), 5.26 – 5.24 (m, 2H, H2^{A}), 5.15 (td, $J_{4D,3D}$ = 9.3 Hz, $J_{4D,5D}$ = 4.7 Hz, 1H, H4^{D}), 5.07 – 4.83 (m, 12H, H2^{B}, H2^{C}), 4.68 (d, *J* = 8.8 Hz, 1H, H1^{A}), 4.23 (d, *J* = 8.0 Hz, 1H, H1^{B}), 4.10 (d, *J* = 7.4 Hz, 1H, H1^{D}), 4.08 – 3.99 (m, 2H, H3^{A}, H5^{A}), 3.88 – 3.82 (m, 3H, H3^{B}, H3^{C}, H5^{D}), 3.76 (td, *J* = 8.8, 4.8 Hz, 1H, H4^{A}), 3.59 – 3.53 (m, 2H, H1^{C}, H4^{C}), 3.33 (dd, *J* = 11.8, 5.1 Hz, 1H, H5^{C}), 3.29 – 3.19 (m, 2H, H4^{B}, H5$^{D'}$), 3.10 – 3.02 (m, 2H, H5$^{A'}$, H5^{B}), 2.67 – 2.56 (m, 2H, H5$^{B'}$,

H5$^{C'}$). **^{13}C NMR** (101 MHz, $CDCl_3$) δ 166.4 (1C, C_q), 166.4 (1C, C_q), 166.4 (1C, C_q), 166.1 (1C, C_q), 165.9 (1C, C_q), 165.8 (1C, C_q), 165.7 (1C, C_q), 165.1 (1C, C_q), 165.1 (3C, C_q), 165.0 (1C, C_q), 164.7 (1C, C_q), 164.4 (1C, C_q), 164.2 (1C, C_q) (*C*O, Bz), 133.9 (2C, Ar*C*H), 133.7 (1C, Ar*C*H), 133.5 (2C, Ar*C*H), 133.4 (1C, Ar*C*H), 133.3(1C, Ar*C*H), 133.2 (2C, Ar*C*H), 133.2(1C, Ar*C*H), 133.1(1C, Ar*C*H), 133.0 (2C, Ar*C*H), 132.9 (2C, Ar*C*H), 132.5 (3C, Ar*C*H), 130.4 (2C, Ar*C*H), 130.2 (2C, Ar*C*H), 130.2 (3C, Ar*C*H), 130.1 (1C, Ar*C*), 130.1 (1C, Ar*C*), 130.0 (3C, Ar*C*H), 129.9 (3C, Ar*C*H), 129.9 (2C, Ar*C*H), 129.9 (2C, Ar*C*H), 129.9 (2C, Ar*C*H), 129.8 (2C, Ar*C*H), 129.8 (2C, Ar*C*H), 129.7 (1C, Ar*C*), 129.7 (1C, Ar*C*), 129.7 (1C, Ar*C*H), 129.6 (1C, Ar*C*), 129.5 (1C, Ar*C*), 129.3 (1C, Ar*C*), 129.2 (1C, Ar*C*), 129.1 (1C, Ar*C*), 129.1 (1C, Ar*C*), 129.0 (1C, Ar*C*), 129.0 (1C, Ar*C*), 129.0 (2C, Ar*C*H), 129.0 (2C, Ar*C*H), 128.9 (2C, Ar*C*H), 128.9 (2C, Ar*C*H), 128.6 (3C, Ar*C*H), 128.6 (2C, Ar*C*H), 128.5 (2C, Ar*C*H), 128.5 (2C, Ar*C*H), 128.4 (3C, Ar*C*H), 128.4 (3C, Ar*C*H), 128.3 (2C, Ar*C*H), 128.3 (2C, Ar*C*H), 128.2 (2C, Ar*C*H), 127.9 (1C, Ar*C*H), 106.1, 105.7, 105.6 (C1^{E}, C1^{F}, C1^{G}), 100.3 (C1^{B}), 100.1 (C1^{C}), 99.9 (C1^{B}), 86.7 (C1^{A}), 82.8, 82.7, 82.5, 82.1, 81.5, 80.8, 78.3, 78.2 (2C), 76.0 (C3^{A}), 75.3, 75.2 (C3^{B}, C3^{C}), 74.5 (C4^{C}), 74.2 (C4^{B}), 73.7(C4^{A}), 73.3 (C2^{B}), 73.0 (C2^{C}), 72.2, 72.0 (C2^{A}, C3^{D}), 71.2 (C2^{D}), 69.8 (C4^{D}), 66.1 (C5^{A}), 64.0, 63.9, 63.8 (C5^{E}, C5^{F}, C5^{G}), 63.1, 63.1, 63.0 (C5^{B}, C5^{C}, C5^{D}). **HRMS:** calcd. for $C_{146}H_{122}O_{43}SNa^+$ *m/z* 2617.6973; found *m/z* 2617.6976.

It has not been possible to assign every peak in the spectra. But the tetra-xylose backbone has been assigned, as well as C1 and C5 for the arabinoses, leaving 18 protons and 9 carbons unassigned corresponding to the unassigned signals that the 3 arabinose units would give rise to. Furthermore, the coupling between the anomeric position in the arabinoses and the corresponding carbon or proton in the xylose tetra-backbone can be seen in HMBC.

2,3,4-Tri-*O*-benzoyl-β-D-xylopyranosyl-(1→4)-[(1→2)-2,3,5-tri-*O*-benzoyl-α-L-arabinofuranosyl]-[(1→3)-2,3,5-tri-*O*-benzoyl-α-L-arabinofuranosyl]-β-D-xylopyranosyl-(1→4)-[(1→3)-2,3,5-tri-*O*-benzoyl-α-L-arabinofuranosyl]-2-*O*-benzoyl-β-D-xylopyranosyl-(1→4)-2,3-di-*O*-benzoyl-D-xylopyranose (119)

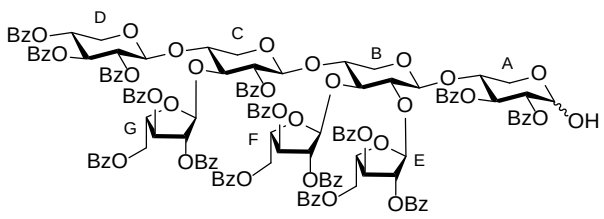


General procedure D: Thioglycoside **109** (0.190 g, 0.073 mmol, 1.0 equiv.). Acetone/H_2O 9:1 (1.1 mL). NBS (0.057 g, 0.32 mmol, 4.4 equiv.), additional NBS added after 30 min (0.033 g, 0.19 mmol, 2.5 equiv.). Reaction time: 60 min.

Eluent for TLC and flash column chromatography (heptane/acetone 3:2). Product was obtained as white crystals (0.169 g, α/β ratio approximately 1:0.3, 92%). $\mathbf{R_f}$(heptane/EtOAc 1:1) 0.24. **^{1}H NMR** (400 MHz, $CDCl_3$) δ 8.22 (d, *J* = 8.1 Hz, 2H, Ar*H*), 8.13 – 8.03 (m, 8H, Ar*H*), 7.99 – 7.94 (m, 8H, Ar*H*), 7.86 – 7.80 (m, 7H, Ar*H*), 7.75 – 7.72 (m, 2H, Ar*H*), 7.69 (d, *J* = 7.4 Hz, 2H, Ar*H*), 7.65 – 7.59 (m, 5H, Ar*H*), 7.56 – 7.25 (m, 31H, Ar*H*, $CHCl_3$), 7.20 – 7.11 (m, 8H, Ar*H*), 7.08 – 7.03 (m, 2H, Ar*H*), 5.79 (t, *J* = 9.5 Hz, 1H, $H3^A$), 5.71 (s, 1H, $H1^G$), 5.64 (t, *J* = 9.6 Hz, 1H, $H3^D$), 5.58 (d, *J* = 2.7 Hz, 1H, $H1^F$), 5.54 (s, 1H), 5.53 – 5.49 (m, 2H, $H1^A$), 5.45 – 5.41 (m, 2H), 5.37 (t, *J* = 4.8 Hz, 2H), 5.29 (s, 0.2H), 5.21 – 5.19 (m, 3H, $H2^D$, $H1^E$), 5.16 – 4.94 (m, 6H, $H2^A$, $H2^B$, $H4^D$), 4.87 – 4.83 (m, 3H), 4.81 – 4.75 (m, 2H), 4.71 (d, *J* = 7.7 Hz, 1H, $H1^A{}_\beta$), 4.38 (d, $J_{1B,2B}$ = 6.5 Hz, 1H, $H1^B$), 4.35 (d, *J* = 6.6 Hz, 0.3H), 3.98 – 3.92 (m, 2H, $H3^B$, $H1^D$), 3.84 – 3.79 (m, 2H, $H1^C$, $H3^C$), 3.76 (t, *J* = 11.0 Hz, 1H, $H5^A$), 3.62 – 3.30 (m, 7H, $H4^A$, $H5^{A'}$, $H4^B$, $H5^B$, $H2^C$, $H4^C$, $H5^D$) 3.23 (t, *J* = 11.2 Hz, 0.4H), 3.10 – 2.96 (m, 3H, $H5^{B'}$, $H5^C$, $H5^{D'}$), 2.49 – 2.44 (m, 1H, $H5^{C'}$), 1.78 (bs, 1H, OH). **^{13}C NMR** (101 MHz, $CDCl_3$) δ 167.1 (1C, C_q), 166.6 (2C, C_q), 166.4 (1C, C_q), 166.3 (1C, C_q), 166.0 (1C, C_q), 165.9 (1C, C_q), 165.8 (1C, C_q), 165.8 (1C, C_q), 165.8 (1C, C_q), 165.8 (1C, C_q), 165.6 (1C, C_q), 165.5 (1C, C_q), 165.5 (1C, C_q), 165.4 (1C, C_q), 165.4 (1C, C_q), 165.0 (1C, C_q), 165.0 (1C, C_q), 164.9 (1C, C_q), 164.9 (1C, C_q), 164.9 (1C, C_q), 164.4 (1C, C_q), 164.4 (1C, C_q) (*C*O, Bz), 134.0 - 128.1 (Ar*C*, Ar*C*H), 106.1 ($C1^F$), 105.7 ($C1^G$), 105.4 ($C1^E$), 105.3 (C_β), 100.2 (C_β), 100.1 ($C1^B$), 99.8 ($C1^D$), 99.6 ($C1^C$), 96.3 (C_β), 90.5 ($C1^A$), 82.0, 82.0, 81.9, 81.8, 81.8 (C_β), 81.7, 81.5, 79.0, 78.6, 78.3, 77.4 ($C2^C$), 76.7 ($C3^C$), 75.4 ($C3^B$), 74.7, 74.5 ($C4^A$, $C4^C$), 74.3 (C_β), 74.2 (C_β), 73.1, 73.1 ($C2^B$,$C4^B$), 72.4, 72.3 ($C2^A$, $C3^D$), 71.1 ($C2^B$), 70.1 ($C3^C$), 69.7 ($C4^D$), 64.3, 64.1, 64.1 ($C5^E$, $C5^F$, $C5^G$), 63.5 (C_β), 63.0, 62.8 ($C5^C$, $C5^D$), 62.0 ($C5^B$), 59.3 ($C5^A$). **HRMS:** calcd. for $C_{140}H_{118}O_{44}S(Na^+)_2$ *m/z* 2548.6777; found *m/z* 2548.6756.

As the α-isomer is the major product this anomer has been partly assigned. However some signals from the β isomer can be seen both in the ^{1}H and ^{13}C NMR spectra, $H1^A{}_\beta$ (4.71 ppm (d, *J* = 7.7 Hz)) and $C1^A{}_\beta$ (96.3 ppm) of the β isomer can be elucidated, as well as others can clearly be seen. It has not been possible to assign every peak for the α-anomer, but the tetra-xylose backbone has been assigned, as well as C1, C5, and H1 for the arabinoses, leaving 15 protons unassigned corresponding to the unassigned signals that the 3 arabinose units would give rise to in the α-anomer. Moreover, the coupling between the anomeric position in the arabinoses and the corresponding carbon or proton in the xylose tetra-backbone can be seen in HMBC.

2,3,4-Tri-*O*-benzoyl-β-D-xylopyranosyl-(1→4)-[(1→ 3)-2,3,5-tri-*O*-benzoyl-α-L-arabinofuranosyl]-2-*O*-benzoyl-β-D-xylopyranosyl-(1→4)- [(1→2)-2,3,5-tri-*O*-benzoyl-α-L-arabinofuranosyl]-[(1→3)-2,3,5-tri-*O*-benzoyl-α-L-arabinofuranosyl]-β-D-xylopyranosyl-(1→4)-2,3-di-*O*-benzoyl-D-xylopyranose (120)

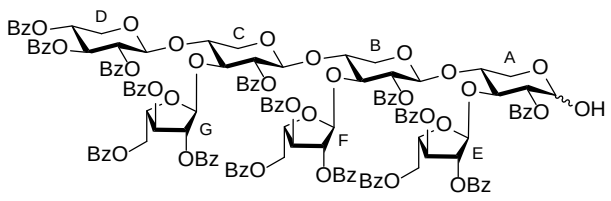


General procedure D: Thioglycoside **114** (0.200 g, 0.077 mmol, 1.0 equiv.). Acetone/H_2O 9:1 (1.23 mL). NBS (0.055 g, 0.31 mmol, 4.0 equiv.), additional NBS added after 30 min (0.030 g, 0.17 mmol, 2.2 equiv.). Reaction time: 60 min. Eluent for TLC and flash column chromatography (heptane/acetone 3:2). Product was obtained as white crystals (0.181 g, α/β ratio approximately 1:0.35, 93%). **R_f**(heptane/EtOAc 1:1) 0.45. **^{1}H NMR** (400 MHz, $CDCl_3$) δ 8.19 – 8.18 (m, 2H, Ar*H*), 8.10 – 7.80 (m, 27H, Ar*H*), 7.78 – 7.74 (m, 1H, Ar*H*), 7.69 – 7.64 (m, 2H, Ar*H*), 7.59 – 7.20 (m, 39H, Ar*H*, $CHCl_3$), 7.18 – 7.07 (m, 4H, Ar*H*), 5.86 – 5.82 (m, 1H, H3^A), 5.79 – 5.78 (m, 1H, H1^F), 5.69 (t, *J* = 9.2 Hz, 1H, H3^D), 5.63 - 5.61 (m, 1H, H1^E), 5.55 – 5.50 (m, 6H), 5.36 – 5.35 (m, 2H, H1^A), 5.26 (s, 1H, H1^G), 5.23 (dd, $J_{2D,3D}$ = 9.3 Hz, $J_{2D,1D}$ = 7.4 Hz, 1H, H2^D), 5.17 (td, $J_{4D,3D}$ = 9.2 Hz, $J_{4D,5D}$ = 5.5 Hz, 1H, H4^D), 5.12 – 5.07 (m, 2H), 5.05 – 4.81 (m, 7H, H2^A, H2^C), 4.76 – 4.64 (m, 2H), 4.57 (d, *J* = 7.3 Hz, 0.33H, H1$^A\beta$), 4.31 (t, $J_{1B,2B}$ = 7.5 Hz, 1H, H1^B), 4.14 (d, $J_{1D,2D}$ = 7.3 Hz, 1H, H1^D), 4.00 (dd, *J* = 11.9 Hz, *J* = 5.3 Hz, 0.40H), 3.93 – 3.82 (m, 6H, H4^A, H5^A, H3^B, H1^C, H3^C, H5^D), 3.74 – 3.69 (m, 1H, H5^A'), 3.66 (dd, $J_{2B,3B}$ = 9.6 Hz, $J_{2B,1B}$ = 6.9 Hz, 1H, H2^B), 3.54 – 3.38 (m, 3H, H4^A, H4^C), 3.36 – 3.24 (m, 2H, H5^B, H5^D'), 3.07 – 3.00 (m, 1H, H5^C), 2.74 – 2.63 (m, 1H, H5^B'), 2.53 (t, *J* = 11.1 Hz, 1H, H5^C'), 1.78 (bs, 1H, OH). **^{13}C NMR** (101 MHz, $CDCl_3$) δ 166.9 (1C, C_q), 166.7 (1C, C_q), 166.5 (1C, C_q), 166.4 (1C, C_q), 166.1 (1C, C_q), 165.9 (1C, C_q), 165.9 (1C, C_q), 165.8 (1C, C_q), 165.8 (1C, C_q), 165.7 (1C, C_q), 165.7 (1C, C_q), 165.6 (1C, C_q), 165.6 (1C, C_q), 165.4 (1C, C_q), 165.4 (1C, C_q), 165.4 (1C, C_q), 165.1 (1C, C_q), 165.0 (1C, C_q), 164.6 (1C, C_q), 164.2 (1C, C_q) (*C*O, Bz), 133.7 (1C, Ar*C*H), 133.6 (1C, Ar*C*H), 133.6 (1C, Ar*C*H), 133.5 (2C, Ar*C*H), 133.4 (1C, Ar*C*H), 133.4 (1C, Ar*C*H), 133.3 (1C, Ar*C*H), 133.2 (1C, Ar*C*H), 133.1 (1C, Ar*C*H), 133.0 (2C, Ar*C*H), 132.9 (1C, Ar*C*H), 132.8 (1C, Ar*C*H), 130.3 (1C, Ar*C*H), 130.2 (C_q, Ar*C*), 130.2 (2C, Ar*C*H), 130.1 (1C, Ar*C*H), 130.0 (2C, Ar*C*H), 130.0 (2C, Ar*C*H), 129.9 (3C, Ar*C*H), 129.9 (3C, Ar*C*H), 129.9 (3C, Ar*C*H), 129.8 (2C, Ar*C*H), 129.8 (3C, Ar*C*H), 129.8 (2C, Ar*C*H), 129.6 (C_q, Ar*C*), 129.6 (C_q, Ar*C*), 129.3 (C_q, Ar*C*), 129.2 (C_q, Ar*C*), 129.2 (C_q, Ar*C*), 129.2 (C_q, Ar*C*), 129.1 (C_q, Ar*C*), 129.1 (C_q, Ar*C*), 129.1 (C_q, Ar*C*), 129.0 (C_q, Ar*C*), 128.9 (C_q, Ar*C*), 128.8 (C_q, Ar*C*), 128.7 (2C, Ar*C*H), 128.7 (2C, Ar*C*H), 128.6 (3C, Ar*C*H), 128.6 (3C, Ar*C*H), 128.5 (3C, Ar*C*H), 128.4 (2C,

Ar*C*H), 128.4 (3C, Ar*C*H), 128.3 (2C, Ar*C*H), 128.2 (3C, Ar*C*H), 107.2 (C_β), 107.1 ($C1^E$), 106.2 ($C1^G$), 105.5 ($C1^F$), 101.3 (C_β), 100.9 ($C1^B$), 99.8 ($C1^D$), 99.6 ($C1^C$), 96.0 (C_β), 90.5 ($C1^A$), 82.7, 82.5 (C_β), 82.3 (C_β), 82.3, 82.0, 81.9, 81.8, 81.8 (C_β), 81.7(C_β), 80.8, 80.3 (C_β) 80.0 ($C2^B$), 78.6, 78.2, 77.9, 77.8 (C_β), 76.0 ($C3^B$/$C3^C$), 75.9 (C_β), 75.7 ($C3^B$/$C3^C$), 74.6 ($C4^C$), 74.2, 74.0 ($C4^A$/$C4^B$), 74.0, 73.8 ($C4^A$/$C4^B$), 73.2 ($C2^C$), 72.4, 72.0 (2C, $C2^A$, $C3^D$), 71.2 ($C2^D$), 70.4 ($C3^A$), 69.7 ($C4^D$), 64.4 (C_β), 64.3, 64.1, 63.9 ($C5^E$, $C5^F$, $C5^G$), 63.1 (C_β), 63.0 ($C5^C$), 62.9 ($C5^D$), 62.4 ($C5^B$), 59.4 ($C5^A$). **HRMS:** calcd. for $C_{140}H_{118}O_{44}S(Na^+)_2$ *m/z* 2548.6777; found *m/z* 2548.6562.

The major α-product has been partly assigned. However, $H1^{A}{}_\beta$ (4.56 ppm (d, *J* = 7.4 Hz)) and $C1^{A}{}_\beta$ (96.0 ppm) of the β can be elucidated from ^{1}H NMR and ^{13}C NMR spectra, furthermore have other clear signal from the β-anomer been indicated. It has not been possible to assign every peak, but the tetra-xylose backbone has been assigned, as well as C1, C5, and H1 for the arabinoses, leaving in the ^{1}H NMR 15 protons unassigned corresponding to the unassigned signals that the 3 arabinose units would give rise to in the α-anomer. Furthermore, the coupling between the anomeric position in the arabinoses and the corresponding carbon or proton in the xylose tetra-backbone can be seen in HMBC. The spectra contain traces of *n*-heptane.

2,3,4-Tri-*O*-benzoyl-β-D-xylopyranosyl-(1→4)-[(1→3)-2,3,5-tri-*O*-benzoyl-α-L-arabinofuranosyl]-2-*O*-benzoyl-β-D-xylopyranosyl-(1→4)-[(1→3)-2,3,5-tri-*O*-benzoyl-α-L-arabinofuranosyl]-2-*O*-benzoyl-β-D-xylopyranosyl-(1→4)-[(1→3)-2,3,5-tri-*O*-benzoyl-α-L-arabinofuranosyl]-2-*O*-benzoyl-D-xylopyranose (121)

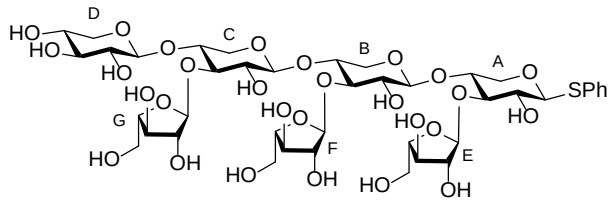


General procedure D: Thioglycoside **118** (0.081 g, 0.031 mmol, 1.0 equiv.). NBS (0.030 g, 0.17 mmol, 5.5 equiv.), additional NBS added after 2 h and 15 min (0.026 g, 0.15 mmol, 4.7 equiv.). Reaction time: 3 h. Eluent for flash column chromatography (heptane/acetone 3:2). The product was obtained as white crystals (0.072 g, α/β ratio 1:0.30 approximately, 92%). **R_f**(heptane/EtOAc 1:1) 0.42. **^{1}H NMR** (400 MHz, $CDCl_3$) δ 8.18 – 7.80 (m, 39H, Ar*H*), 7.64 – 7.19 (m, 58.5H, Ar*H*, $CHCl_3$), 5.78 (t, *J* = 8.9, 0.2H) 5.71 (t, *J* = 9.3 Hz, 1H, $H3^D$), 5.53 – 5.49 (m, 4H, $H1^E$/$H1^F$), 5.46 – 5.42 (m, 2H), 5.39 – 5.35 (m, 4H, $H1^A$), 5.30 – 5.28 (m, 2H, $H2^D$, $H1^G$), 5.26 (s, 1H, $H1^E$/$H1^F$), 5.16 (td, *J* = 9.3 Hz, *J* = 5.5 Hz, 1H, $H4^D$), 5.09 – 4.84 (m, 15H, $H2^A$, $H2^B$, $H2^C$), 4.48 (d, *J* = 7.8 Hz, 0.3H, $H1^{A}{}_\beta$), 4.36 (t, *J* = 9.3 Hz, 1H, $H3^A$), 4.25 (dd, *J* = 8.0 Hz, *J* = 3.7 Hz, 1H, $H1^B$),

4.11 – 4.05 (m, 2H, H1^{D}), 3.91 – 3.81 (m, 4H, H3^{B}, H3^{C}, H5^{D}), 3.76 (td, *J* = 9.5 Hz, *J* = 5.0 Hz, 1H, H4^{A}), 3.58 – 3.46 (m, 4H, H5^{A}, H5$^{A'}$, H1^{C}, H4^{C}), 3.37 – 3.29 (m, 2H, H4^{B}, H5^{C}), 3.29 – 3.24 (m, 1H, H5$^{D'}$), 3.12 (dd, *J* = 11.9 Hz, *J* = 5.6 Hz, 1H, H5^{B}), 2.99 – 2.94 (m, 0.3H), 2.66 – 2.59 (m, 2H, H5$^{B'}$, H5$^{A'}$,), 1.61 (bs, 4H, OH). **^{13}C NMR** (101 MHz, $CDCl_3$) δ 167.0 ($C_β$), 166.4 (1C, C_q), 166.4 (1C, C_q), 166.4 (1C, C_q), 166.1 (1C, C_q), 165.9 (1C, C_q), 165.8 (1C, C_q), 165.8 ($C_β$), 165.8 (1C, C_q), 165.6 (1C, C_q), 165.2 (1C, C_q), 165.1 (1C, C_q), 165.1 (1C, C_q), 165.0 (1C, C_q), 164.8 (1C, C_q), 164.5 ($C_β$), 164.4 (1C, C_q), 164.2 (1C, C_q) (*C*O, Bz), 133.9 (1C, Ar*C*H), 133.9 (1C, Ar*C*H), 133.7 (1C, Ar*C*H), 133.7 (1C, Ar*C*H), 133.5 (2C, Ar*C*H), 133.4 (1C, Ar*C*H), 133.4 (1C, Ar*C*H), 133.3 (1C, Ar*C*H), 133.2 (1C, Ar*C*H), 133.2 (1C, Ar*C*H), 133.2 (1C, Ar*C*H), 133.0 (1C, Ar*C*H), 133.0 (2C, Ar*C*H), 132.9 (1C, Ar*C*H), 130.4 (1C, Ar*C*H), 130.4 (1C, Ar*C*H), 130.2 (1C, Ar*C*H), 130.2 (2C, Ar*C*H), 130.1 (1C, Ar*C*H), 130.1 (1C, Ar*C*H), 130.0 (4C, Ar*C*H), 129.9 (2C, Ar*C*H), 129.9 (2C, Ar*C*H), 129.8 (2C, Ar*C*H), 129.8 (2C, Ar*C*H), 129.8 (3C, Ar*C*H), 129.7 (1C, Ar*C*H), 129.6 (C_q, Ar*C*), 129.5 (C_q, Ar*C*), 129.5 (C_q, Ar*C*), 129.5 (C_q, Ar*C*), 129.3 (C_q, Ar*C*), 129.2 (C_q, Ar*C*), 129.2 (C_q, Ar*C*), 129.1 (C_q, Ar*C*), 129.1 (C_q, Ar*C*), 129.0 (2C, C_q, Ar*C*), 129.0 (2C, C_q, Ar*C*), 129.0 (C_q, Ar*C*), 128.9 (1C, Ar*C*H), 128.9 (1C, Ar*C*H), 128.9 (1C, Ar*C*H), 128.7 (1C, Ar*C*H), 128.6 (3C, Ar*C*H), 128.5 (1C, Ar*C*H), 128.5 (1C, Ar*C*H), 128.4 (2C, Ar*C*H), 128.4 (2C, Ar*C*H), 128.4 (2C, Ar*C*H), 128.3 (1C, Ar*C*H), 128.3 (1C, Ar*C*H), 128.3 (1C, Ar*C*H), 128.3 (1C, Ar*C*H), 106.1 (C1^{G}), 105.7, 105.6 (C1^{E}, C1^{F}), 100.4 ($C_β$), 100.3 (C1^{B}), 100.0 (C1^{C}), 100.0 ($C_β$), 99.9 (C1^{D}), 96.4 ($C_β$) 90.5 (C1^{A}), 82.8, 82.7, 82.7, 81.8 ($C_β$), 81.7, 81.5, 80.8, 78.3, 78.3, 78.2, 78.1 ($C_β$), 76.3 ($C_β$), 75.4 ($C_β$), 75.3, 73.3 (C3^{B}, C3^{C}), 74.7 ($C_β$), 74.5, 74.3, 74.3 (C4^{A}, C4^{B}, C4^{C}), 74.2 ($C_β$), 74.2 ($C_β$), 74.1 ($C_β$), 73.4 (C2^{B}), 73.0 (C2^{C}), 72.1, 72.0 (C2^{A}, C3^{D}), 71.2 (C2^{D}), 69.8 (C4^{D}), 64.0, 63.9, 63.9, 63.8 (C5^{E}, C5^{F}, C5^{G}), 63.3 ($C_β$), 63.1, 63.0, 63.0 (C5^{B}, C5^{C}, C5^{D}), 59.3 (C5^{A}). **HRMS:** calcd. for $C_{140}H_{118}O_{44}SNa^+$ *m/z* 2525.6888; found *m/z* 2525.6876.

It has not been possible to assign every peak, but the xylan-backbone as well as C1, C5, and H1 for the arabinose units of the α-anomer have been assigned. Also a few signals from the β-anomer can be elucidated i.e. H1$^{A}{}_{β}$ (4.48 ppm (d, *J* = 7.8 Hz)) and C1$^{A}{}_{β}$ (96.3 ppm) from ^{1}H NMR and HSQC, as well as other signals have been indicated. The spectre contain traces of *n*-heptane.

Phenyl β-D-xylopyranosyl-(1→4)-[(1→3)-α-L-arabinofuranosyl]-β-D-xylopyranosyl-(1→4)-[(1→3)-α-L-arabinofuranosyl]-β-D-xylopyranosyl-(1→4)-[(1→3)-α-L-arabinofuranosyl]-1-thio-β-D-xylopyranose (124)



Thioglycoside **118** (0.027 g, 0.010 mmol, 1.0 equiv.) was dissolved in a 1:1 mixture of dry CH_2Cl_2 and MeOH (1.04 mL). Freshly prepared 1 M NaOMe was added until basic. After 4 h TLC(heptane/acetone 1:1) confirmed that all starting material had reacted, and an additional TLC (EtOAc/MeOH, H_2O, AcOH 5.5:3.5:0.8:0.2) showed the formation of a product. The reaction mixture was neutralised to pH 7 by addition of 0.1 M HCl. The mixture was washed with Et_2O and EtOAc and the aqueous phase was freeze dried (0.016 mg, 151 %). $\mathbf{R_f}$(EtOAc/MeOH, H_2O, AcOH 5.5:3.5:0.8:0.2) 0.31. **^{1}H NMR** (800 MHz, D_2O) δ 7.62 – 7.59 (m, 2H, ArH_{ortho}), 7.47 – 7.43 (m, 3H, ArH_{meta}, ArH_{para}), 5.44 (s, 1H, H1^{Araf}), 5.42 (s, 1H, H1^{Araf}), 5.41 (s, 1H, H1^{Araf}), 4.83 (d, *J* = 9.6 Hz, 1H, H1^{A}), 4.51 – 4.48 (m, 2H, H1^{B}, H1^{C}), 4.45 (d, *J* = 7.8 Hz, 1H, H1^{D}), 4.32 – 4.27 (m, 3H, , H4^{Araf}), 4.20 – 4.16 (m, 4H, H5^{A}, H2^{Araf}), 4.08 (ddd, *J* = 11.5, 4.7, 2.7 Hz, 2H, H5^{B}, H5^{C}), 3.95 – 3.91 (m, 4H, H5^{D}, H3^{Araf}), 3.84 – 3.79 (m, 7H, H3^{A}, H4^{A}, H4^{B}, H4^{C}, H5^{Araf}), 3.78 – 3.72 (m, 5H, H3^{B}, H3^{C}, H5'^{Araf}), 3.62 (td, *J* = 9.8, 5.5 Hz, 1H, H4^{D}), 3.54 (t, *J* = 8.8 Hz, 1H, H2^{A}), 3.48 – 3.41 (m, 4H, H5'^{A}, H2^{B}, H2^{C}, H3^{D}), 3.38 (t, *J* = 11.1 Hz, 1H, H5'^{B}, H5'^{C}), 3.29 (t, *J* = 11.2 Hz, 1H, H5'^{D}), 3.28 – 3.24 (m, 1H, H2^{D}). **^{13}C NMR** (101 MHz, $CDCl_3$) δ 132.4 (2C, ArC_{ortho}H), 131.1 (1C, ArC_{ipso}) 129.3 (2C, ArC_{meta}H), 128.4 (1C, ArC_{para}H), 107.7, 107.6, 107.5 (C1^{Araf}), 101.4 (C1^{D}), 101.4, 101.2 (C1^{B}, C1^{C}), 87.8 (C1^{A}), 84.7, 84.7, 84.7 (C4^{Araf}), 80.7, 80.7, 80.6 (C2^{Araf}), 78.7 (C3^{A}), 77.2 (3C, C3^{Araf}), 77.1 (2C, C3^{B}, C3^{C}), 75.5 (C3^{D}), 73.6, 73.6, 73.6 (C4^{A}, C4^{B}, C4^{C}), 73.5, 73.5 (C2^{B}, C2^{C}), 72.9 (C2^{D}), 71.9 (C2^{A}), 69.1 (C4^{D}), 66.2 (C5^{A}), 65.0 (C5^{D}), 62.7. 62.7 (C5^{B}, C5^{C}), 61.2 (3C, C5^{Araf}). **MS:** calcd. for $C_{41}H_{61}O_{28}S$ (negative mode-H) *m/z* 1033.31; found *m/z* 1033.38.

3. Synthesis of glucuronoxylan fragments

The contents of this chapter represent the results obtained for the chemical synthesis of glucuronoxylan fragments and both results from former Ph.D. students in the group and own results will be described and discussed. The project was initiated by former Ph.D. student Clotilde d'Errico,[150] with additional contributions from former Ph.D. student Maximilian Felix Böhm.[105] The author of the present dissertation received the project in the end phase, but due to lack of material most of the previous work was replicated. The first part of this chapter will demonstrate the results obtained by C. d'Errico[150] and M. F. Böhm.[105]. The second part will illustrate the synthetic steps done for the formation of two glucuronoxylan target structures produced for this thesis.

3.1 Aim of the project

In line with the project regarding the synthesis of arabinoxylans fragments, two pure and well-defined glucuronoxylan fragments for elucidation of hemicellulosic enzymes was pursued (Figure 3-1). Like the AX fragments also the GX fragments consists of a common β-(1→4)-linked xylose backbone differing in the type of glucuronic acid substitution. Consequently, the same synthetic protocol for synthesis of the xylose backbone was utilised (see chapter 2) with inclusion of one new building block to be able to achieve subsequent installation of the glucuronic acid residue as an α-(1→2)-linkage. It was furthermore decided to apply a pre-

glycosylation oxidation protocol for formation of the fragments, hence two glucuronic acid building blocks were needed.

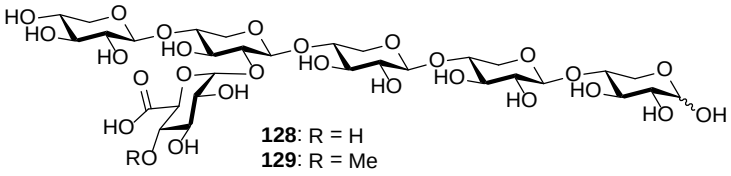


Figure 3-1: Structures of the two GX target structures pursued for this project.

The specific structures were chosen since the backbone should be long enough for good enzyme activity and the position of the substitution will allow for verification of the substrate specificity of xylanases. E.g. would a GH30 be able to cleave the xylose unit at the non-reducing end if the enzyme was incubated with target structure **129** with a MeGlcA substituent, however no cleavage should be observed if target structure **128** was used as the substrate (Figure 3-2). Furthermore, the targets can be used to investigate the substrate specificity of α-glucuronidases as belonging either to GH115 or GH67, as a GH115 would be able to cleave the targets directly, but a GH67 only be able to hydrolyse the glucuronic acid moiety after the action of a GH30 or a GH10.

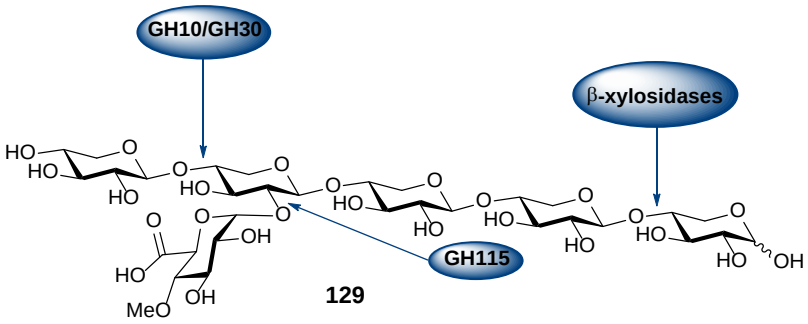


Figure 3-2: Example of possible cleavage sites by enzymatic degradation with glycosyl hydrolases using target **129** as the substrate.

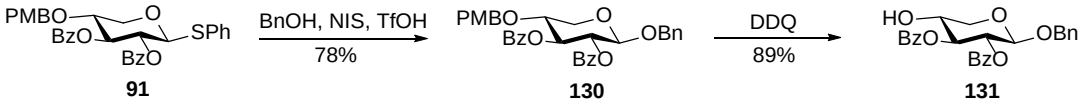
3.2 Previous results for the synthesis of GX fragments

The research presented in this section (section 3.2) represents work done by C. d'Errico[150] *and M. F. Böhm.*[105]

3.2.1 Monomeric xylose building blocks

Following the synthetic strategy already established for the synthesis of AX fragments, the monomeric perbenzoylated building block **80**, as well as the xylose building block for linear elongation **81** can be directly used for the synthesis of the backbone. However, the synthesis of two other monomeric xylose building blocks were needed.

The glycosyl acceptor **131** was prepared from **91** by glycosylation with BnOH in the presence of NIS and TfOH followed by oxidative cleavage of the PMB group as described above (Scheme 3-1).[150] This building block is to be used as the reducing end xylose unit of the pentasaccharide backbone to avoid any collateral reactivity in the last coupling to the GlcA branch.



Scheme 3-1: Synthesis of the reducing end xylose building block **131**.[150]

The other building block needed should allow for the later attachment of a glucuronic acid substituent at the *C*-2 position. Hence, a synthetic protocol for a building block with the temporary levulinoyl protection group attracted at *O*-2 position and the permanent benzoyl group at *O*-3 position was needed (Figure 3-3). However, development of this protocol proved nontrivial.

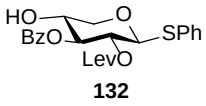


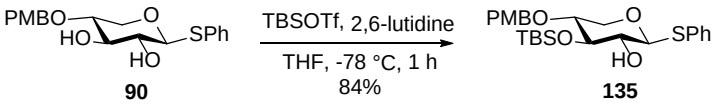
Figure 3-3: Xylose building block **132** with the temporary Lev group at the 2-position.

It was known from the synthesis of the building block **84**, that a regioselective benzoylation of **90** using a phase-transfer catalyst primarily gave the *O*-2 benzoylated product contrary to the generally accepted reactivity of the xylose hydroxyl groups (Scheme 2-11). Attempts to apply the same condition for the regioselective introduction of the Lev groups gave poor conversion and selectivity for the *O*-3 Lev compound (Scheme 3-2).[150] An explanation for the observed reversed reactivity could be that the Lev group is less prone to migrate than the benzoyl group.[151]

Scheme 3-2: Attempted regioselective levulination of **90**.[150]

After several trials for further optimisation of the protocol and trials with other regioselective methods for either benzoylation or levulination of **90** (results not shown),[150] C. d'Errico decided on another approach. By inspiration from an article of Nicolaou *et al.* from 2000,[152] it was found

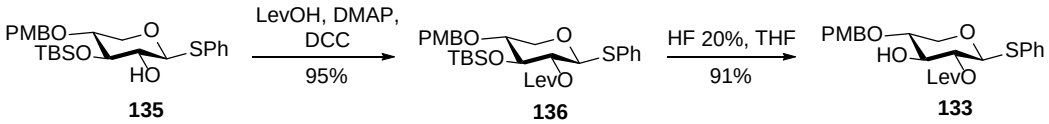
that a *tert*-butyldimethylsilyl (TBS) group could be installed as a temporary protection group at *O*-3 in a solvent dependent matter with remarkable regioselectivity. When compound **90** was reacted with *tert*-butyldimethylsilyl triflate (TBSOTf) and 2,6-lutidine at -78 °C in tetrahydrofuran (THF) the *O*-3 substituted product could be obtained in 84% yield (Scheme 3-3), but if dichloromethane was used as the solvent a mixture of the 2-TBS and the disubstituted product was obtained (40 and 47% yield, respectively).[150]



Scheme 3-3: Regioselective introduction of the temporary TBS group.

The remarkable selectivity might be explained by silyl migration under the basic and kinetic reaction conditions employed.[153–155] It could be assumed that the presence of a base and the polar but aprotic solvent THF allows for migration of the silyl group from the kinetically favoured 2-position to the thermodynamically favoured 3-position. On the contrary, the apolar solvent dichloromethane gave rise to the typical kinetic selectivity.

Then the Lev group could be installed at the *O*-2 position under the standard condition previously established (Scheme 3-4).[150] Subsequent removal of the TBS group proved to be nontrivial due to the presence of both the base-sensitive Lev group and the acid-sensitive PMB group. Several different fluoride sources were tried as silyl ethers are well known to be removable by addition of a fluoride ion due to the high affinity of fluoride to silicon ($\sigma_{Si\text{-}O}$ = 368 kJ/mol, $\sigma_{Si\text{-}F}$ = 582 kJ/mol).[156] When mild conditions with tetrabutylammonium fluoride and a catalytic amount of acetic acid was used, the main product was the diol **90**. Hence, the applied conditions still hydrolysed the Lev group despite the presence of acid. Using a stoichiometric amount of acid no reaction had occurred after 48 h, and the same was observed for a reaction with a HF-pyridine complex (results not shown).[150] Finally, it was found that the TBS group could be selectively removed by dissolving **136** in THF in low concentration followed by addition of a 20% solution of HF in water (Scheme 3-4).[150]



Scheme 3-4: Levulination of **135**, followed by deprotection of the TBS group providing **133**.[150]

After successful removal of the TBS group, a benzoyl group was installed at the 3-position (77%) and the PMB group removed (92%) both under standard conditions providing the building block **132** (37% overall yield over five steps from compound **90**).[150]

3.2.2 Monomeric glucuronic acid building blocks

Two different glucuronic acid donors were needed for construction of the target structures. As the glycosylations with the donors should provide α-(1,2)-linkages, the donors needed a non-participating group at *O*-2 (see Section 1.6). The protection group of choice was the benzyl group, which at the same time increases the reactivity of the donor as it is an arming protection group. The methyl ester was chosen as protection for the carboxylic acid to reduce the polarity and limit side reactions. Based on previous utilisations of glucuronic acid thioglycoside donors by Oscarson's group[76,157] the thioethyl group was the donor functionality of choice. The two donors are displayed in Figure 3-4. They differed in the substituent of the 4-position, which was either a benzyl (**137**) or a methyl ether (**138**) representing either a GlcA or a MeGlcA substituent, respectively.

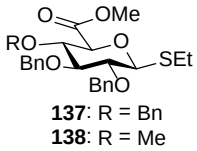
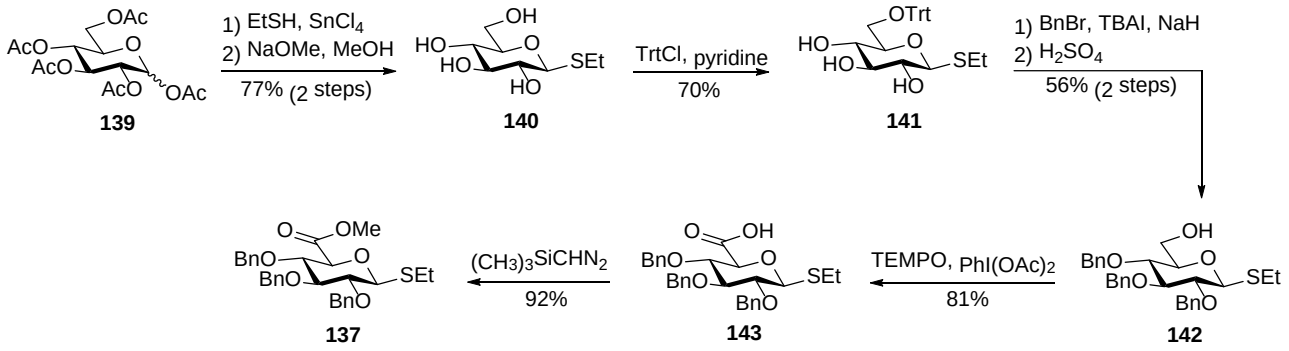


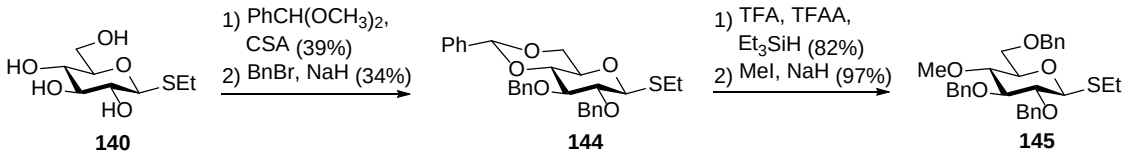
Figure 3-4: Target glucuronic acid donors.

Building block **137** was synthesised by C. d'Errico following a fairly straightforward protocol.[150] Ethyl 1-thio-β-D-glycopyranoside (**140**) was synthesised by condensation of peracetylated glucose **139** and thioethanol in the presence of $SnCl_4$ as a Lewis promotor according to a literature protocol[158] (Scheme 3-5). The tetraol **140** was then formed by removal of the acetates under Zemplén conditions. By reaction with the bulky trityl chloride the primary alcohol could be selectively protected using traditional conditions. This made benzylation of the secondary alcohols possible through a Williamson's ether synthesis, followed by hydrolysis of the trityl ether under acidic conditions without purifying the intermediate. The two last synthetic steps before completing the synthesis of **137** was the oxidation of the primary alcohol to the carboxylic acid and subsequent methylation, this was achieved by reaction with 2,2,6,6-tetramethyl-1-piperidinyloxy (TEMPO) and an excess of (diacetoxyiodo)benzene[159,160] and (trimethylsilyl)diazomethane (2 M in hexane), respectively (the mechanism of the latter will be discussed *vide infra*).



Scheme 3-5: Synthesis of glucuronic acid donor **137** as reported by C. d'Errico.[150]

For the glucuronic acid building block **138** the first strategy investigated by C. d'Errico involved the introduction of a benzylidene acetal at the *O*-4 and *O*-6 positions of **140** under standard reaction conditions through reaction with benzaldehyde dimethylacetal under acidic conditions. The reaction proceeded in a relatively low yield of 39% after a reaction time of 24 h, but some starting material was also recovered (Scheme 3-6). The next transformation was the benzyl protection of the remaining hydroxyl groups providing **144**, followed by reductive ring-opening of the benzylidene acetal by reaction with trifluoroacidic acid (TFA), trifluoroacetic anhydride (TFAA), and triethylsilane, to give the free 4-OH glucoside. Methylation of the free hydroxyl group via Williamson's ether synthesis with methyl iodide as the methylating agent smoothly provided the fully protected glucoside **145**.[150]

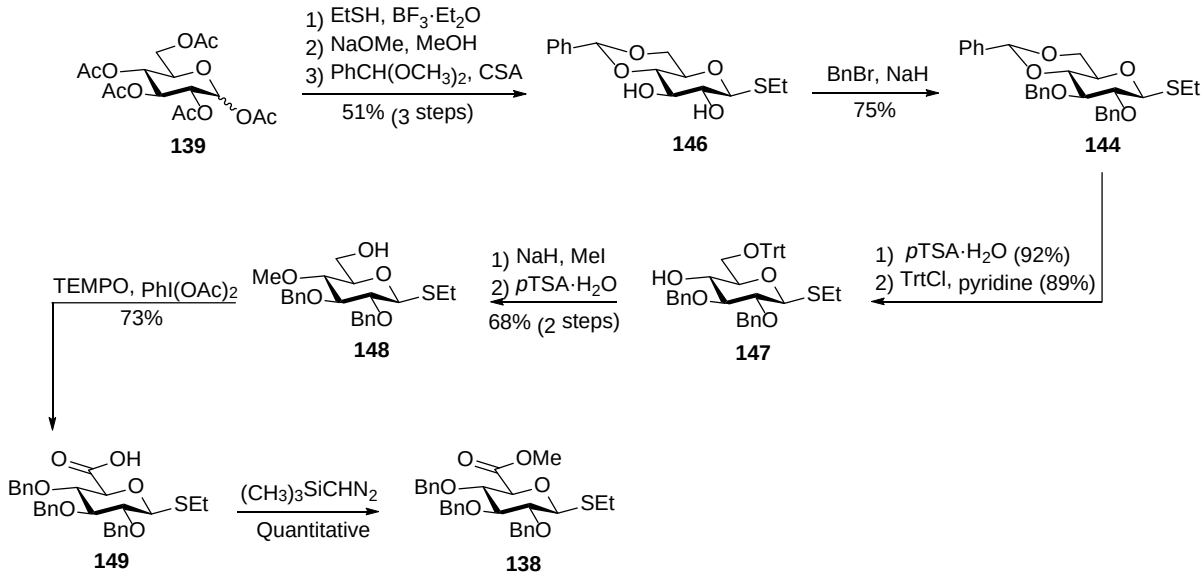


Scheme 3-6: Synthetic steps achieved by C. d'Errico for the synthesis of glucuronic acid building block **138**.[150]

In the following, it was envisioned to liberate the primary hydroxyl group of **145** by selective hydrolysis and then subsequent oxidation to the carboxylic acid and methylation to the ester using the same conditions as established for **137**. The selective hydrolysis of the benzyl ether with DDQ has been reported for a few studies,[161,162] but when applying the conditions analogous to the PMB deprotection, the main product was the compound with a free 3-OH. The same outcome was obtained using DDQ under anhydrous conditions and UV irradiation.[150] A protocol developed by Fraser-Reid and co-workers using $FeCl_3$ under strict anhydrous conditions[163] also failed to provide the expected primary alcohol, instead the 1,6-anhydro

glucopyranose was obtained. Hence C. d'Errico concluded that to bypass the encountered difficulties another synthetic strategy was needed.[150]

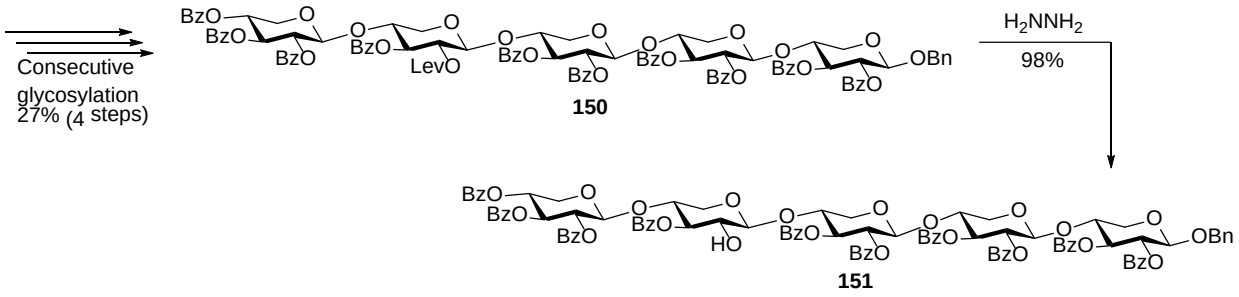
The synthesis of glucuronic acid donor **138** was then synthesised in 10 steps and 16% overall yield from compound **139** by M. F. Böhm[105] following the same transformations as a protocol published by Oscarson and Svahnberg in 2001[76] obtaining the same product (the individual synthetic steps will be discussed *vide infra*) (Scheme 3-7).



Scheme 3-7: Synthesis of donor **138** as reported by F. M. Böhm[105] following the protocol of Oscarson and Svahnberg.[76]

3.2.3 Synthesis of GX target structures

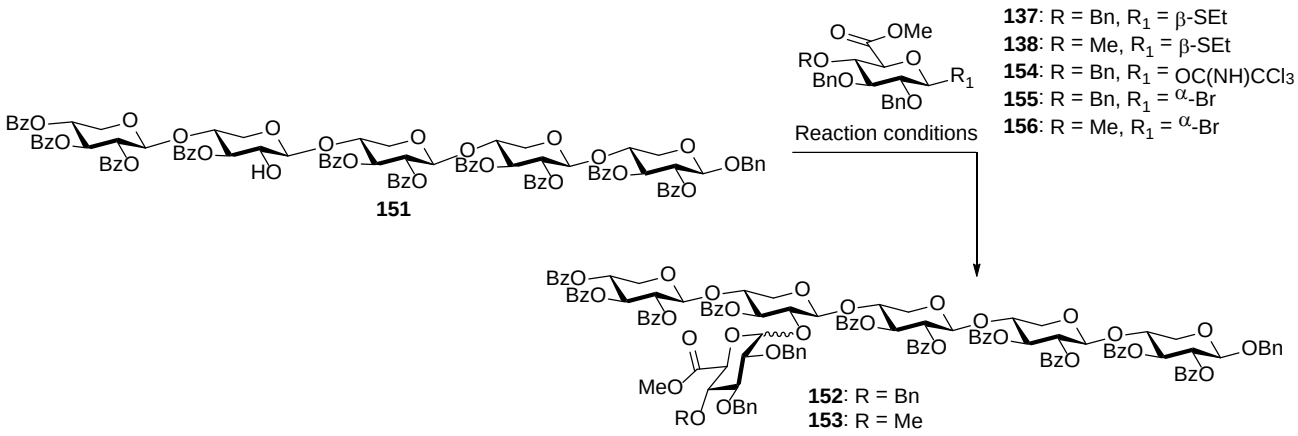
The pentasaccharide backbone of xylose (**150**) was synthesised by C. d'Errico in four consecutive glycosylations in an overall yield of 27%[150] following the same protocol as established for the synthesis of the arabinoxylan fragments.[105] The following deprotection of the Lev group under the previously established condition provided the deprotected pentasaccharide acceptor **151** although flash chromatography of the product was needed (Scheme 3-8).[105]



Scheme 3-8: Assembly of pentasaccharide **151** following the pre-established glycosylation protocol followed by Lev deprotection.[150]

With the acceptor pentasaccharide **151** in hand several conditions for the glycosylation with the donors **137** and **138** or derivatives hereof were tried (Table 3-1).

Table 3-1: Investigated reaction conditions for the formation of the hexasaccharides **152** and **153** by M. F. Böhm and C. d'Errico.



Entry	Donor	Promotor	Reaction conditions	Result	Reference
1	**138** (1.5 equiv.)	DMTST (4.0 equiv.)	0 °C, Et_2O, 20 h	Never went to completion, decomposition of donor	M. F. Böhm[105]
2	**138** (1.5 equiv.)	NIS (1.6 equiv.)/ TESOTf (0.2 equiv.)	-10 °C, CH_2Cl_2/Et_2O 1:1, 1 h	No reaction, decomposition of donor	M. F. Böhm[105]
3	**138** (3.0 equiv.)	*p*NO_2-PhSCl (1.2 equiv.), AgOTf (2.0 equiv.)	-40 °C, CH_2Cl_2, 20 min	**153** (61%), α/β 1:0.9	C. d'Errico[a]
4	**154** (1.5 equiv.)	TMSOTf (0.1 equiv.)	0 °C, dioxane/Et_2O 1:1, 2 h	**152** (29%), α/β 1:1	C. d'Errico[a]

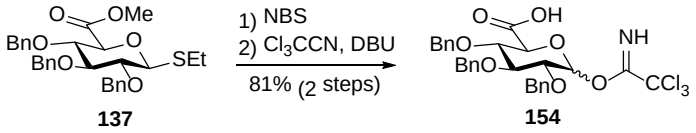
5	**156** (1.5 equiv.)[b]	AgOTf (1.8 equiv.)	-30 °C, CH_2Cl_2	**153** (42%), α/β 1.6:1	M. F. Böhm[105]
6	**156** (1.5 equiv.)[b]	$AgClO_4$ (1.8 equiv.)	-30 °C, CH_2Cl_2/Et_2O 9:1	**153** (64%), α/β 2.5:1 **151** was recovered in 15%	M. F. Böhm[105]
7	**156** (1.5 equiv.)[b]	$AgClO_4$ (1.2 equiv.)	-30 °C, CH_2Cl_2/Et_2O 98:2, 2 h	**153** (55%), α/β 3:1 **151** recovered	C. d'Errico[a]
8	**155** (1.5 equiv.)[b]	$AgClO_4$ (1.2 equiv.)	-30 °C, CH_2Cl_2/Et_2O 98:2, 2 h	No reaction	C. d'Errico[a]
9	**137** (1.5 equiv.)	TBAB (5.0 equiv.)	r.t., titration with Br_2 (1 M in CH_2Cl_2), CH_2Cl_2/Et_2O 98:2, 48 h	No reaction	C. d'Errico[a]
10	**137** (1.5 equiv.)	$AgClO_4$ (1.2 equiv.)	-30 °C, titration with Br_2 (0.2 M in CH_2Cl_2), CH_2Cl_2/ Et_2O 3:1, 1.5 h	**152** (39%), α/β 5:1	C. d'Errico[a]
11	**137** (4.0 equiv.)	$AgClO_4$ (4.8 equiv.)	-30 °C, titration with Br_2 (1 M in CH_2Cl_2), CH_2Cl_2/ Et_2O 3.5:1, 1.5 h	**152** (59%), α/β 3:1 **151** was recovered in 24%	C. d'Errico[a]
12	**138** (4.0 equiv.)	$AgClO_4$ (4.8 equiv.)	-30 °C, titration with Br_2 (1 M in CH_2Cl_2), CH_2Cl_2/ Et_2O 4:1, 2 h	**153** (56%), α/β 2:1 **151** was recovered in 34%	C. d'Errico[a]

[a]Unpublished results, produced during a short Postdoc for C. d'Errico. [b]Bromide formed by reaction with Br_2 (4.0 equiv., 0.4 M in CH_2Cl_2), then the mixture was quenched, washed, and solvent removed after which the crude was used as donor.

The first conditions investigated were based on the work by Oscarson and Svahnberg where excellent α-selectivity is obtained for similar substrates[76] using DMTST in diethyl ether (Entry 1). Under these conditions some product formation was observed, however it was the donor (**138**) and not the acceptor (**151**) that was fully consumed, even after addition of extra donor. An explanation might be the low solubility of the acceptor **151** in diethyl ether. When the NIS/triethylsilyl trifluoromethanesulfonate (TESOTf) promotor system in dichloromethane/diethyl ether (1:1) was applied, the donor decomposed after one hour at -10 °C (Entry 2). Application of the conditions for the optimised promotor system for the xylose building blocks resulted in formation of the product **153** in 61% yield, but with poor stereoselectivity (α/β 1:0.9) (Entry 3).

As poor results were obtained for the thioethyl donor **138** it was decided to try other donor functionalities. The troubles reflect the low reactivity of uronic acid donors, which is the

consequence of the chosen pre-glycosylation oxidation approach. The first donor functionality tried was the trichloroacetimidate, which was formed through removal of the thioethyl moiety by reaction with NBS followed by reaction with trichloroacetonitrile and base (Scheme 3-9).



Scheme 3-9: Formation of trichloroacetimidate donor **154** from **137**.

When acetimidate **154** was utilised as the donor, a low yield and a poor stereoselectivity was observed (Entry 4). Hence, attention was turned towards using the anomeric α-bromides (**155** and **156**) as a donors, where the bromide would increase the reactivity of the donor. The first trial was with the α-bromide **156**, which was formed from the thioethyl donor **138** by reaction with bromine, and the reaction was subsequently quenched, washed, the solvent evaporated, and the crude **156** used directly for the glycosylation promoted by silver triflate (Entry 5). These conditions provided the product in moderate yield in a clean reaction and showed improvement in the stereoselectivity. To further shift the ratio towards the α-product, silver perchlorate was tried as the promotor and the solvent composition was changed to dichloromethane/diethyl ether 9:1 ($c_{acceptor}$ 0.15 M) (Entry 6). As expected, it was found that these conditions lead to both a better yield and a better α-stereoselectivity and some of the acceptor could be reisolated. Similar conditions were tried by C. d'Errico for both bromide donors but with approximately one third the concentration ($c_{acceptor}$ 0.055 M Entry 7 and $c_{acceptor}$ 0.062 M Entry 8) and the dichloromethane/diethyl ether ratio adjusted to 98:2 for better solvation. It was found that for donor **156** the reaction proceeded in good yield and better stereoselectivity, but when donor **155** was used, the reaction surprisingly did not proceed.

As the reaction with bromide donor **155** and silver perchlorate did not seem to proceed, the same donor was formed *in situ* from thioethyl donor **137** and glycosylation tried under the promotion of tetrabutylammonium bromide (TBAB) inspired by a protocol developed by Lemieux *et al.*[164] (Entry 9). Also under these conditions, the reaction did not proceed even after an extended reaction time of 48 hours.

As the bromide formed from thioethyl donor **138** had given the best results, the bromide **155** was tried again as the donor for the glycosylation, but this time the bromide was formed *in situ* from **137** by titration with bromine (0.2 M) until a yellow colour of the reaction mixture

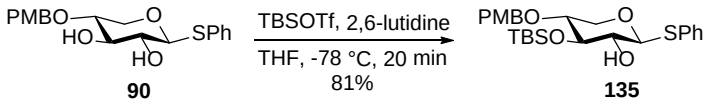
persisted (Entry 10). Furthermore, the solvent composition was adjusted raising the diethyl ether content (dichloromethane/diethyl ether 3:1) and the concentration of the reaction mixture lowered even further ($c_{acceptor}$ 0.023 M). Under these conditions, it was seen that product **152** was produced in moderate yield but with a better stereoselectivity (α/β 5:1). Inspired by these results it was attempted to raise the yield by addition of more donor and promotor (Entry 11). This lead to formation of **152** in a higher yield, but with a bit lower stereoselectivity after two hours of reaction time. Furthermore, the acceptor **151** could be reisolated. When the same conditions were applied for donor **138** comparative results were achieved (Entry 12).

The α/β mixtures of **152** and **153** produced in the described glycosylations were obtained as inseparable mixtures.

3.3 Synthesis of a monomeric xylose building block

The results presented in this and following sections of this chapter represent research done by the author for this thesis.

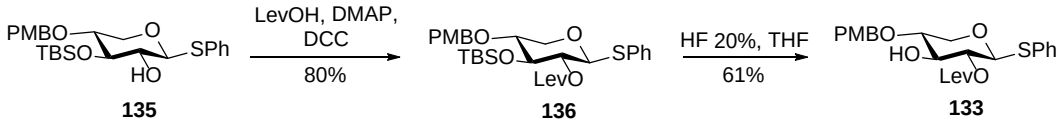
As mentioned above the author of this dissertation entered this project at late stage in the synthesis, but as more material was needed, several of the syntheses shown *vide supra* were reproduced. The synthesis of xylose building block **132** was performed starting from diol **90**. The TBS group was selectively introduced at the 3-position following the conditions established above (Scheme 3-10).



Scheme 3-10: Introduction of a TBS group at the 3-position of **90**.

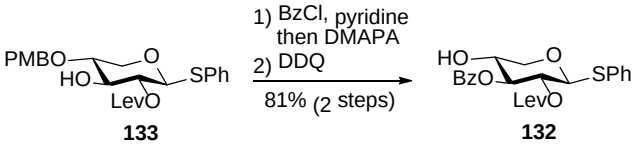
It was found that the reaction time could be reduced to 20 min and while still obtaining a comparable yield of **135**. It is worth noticing that the presence of the bulky TBS group influenced the conformation of the compound, as these groups are known to do.[165] The presence of the TBS group at the 3-position generated a distortion from the common $^4C_1$ chair to somewhere between the chair and a half-chair. This can be seen by the visibly lowered coupling constant between H1 and H2 ($J_{1,2}$ = 3.0 Hz). A similar value can also be observed for H2 and H3 ($J_{3,2}$ = 4.4 Hz). The explanation for this conformational change is that the molecule in this way reduced the strain between the two bulky vicinal groups TBS and PMB.

Subsequently the Lev group was introduced on the 2-position, but unfortunately the following TBS deprotection did not proceed with as much success as observed for C. d'Errico (Scheme 3-11). The TBS cleavage was repeated several times without improving the yield. It is speculated that the moderate yield could be due to simultaneous hydrolysis of the acid-sensitive PMB group, as PMBOH was isolated in an amount corresponding to hydrolysis of 5% of the starting material for one reaction. This further confirms that for compound **136** removal of the TBS moiety is a tricky reaction, but due to time constraints, the conditions were not further optimised.



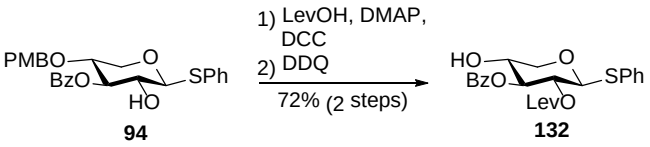
Scheme 3-11: Formation of **133** by levulination of **135** and deprotection of the TBS group.

Introduction of the benzoyl group was attempted under the same conditions (BzCl, Et_3N, DMAP) as for synthesis of compound **91**, as to avoid flash chromatography purification after the reaction by quenching with DMAPA.[116] However, when these conditions were used the reaction was slow and the starting material was left in a considerable amount even after 24 h reaction time. Hence, the benzoyl group was introduced in a similar fashion as for compound **80** and **96** (BzCl in pyridine). The reaction medium was again quenched by DMAPA, providing a clean compound after washing. Consequently, the crude product was directly subjected to oxidative cleavage of the PMB group providing **132** (Scheme 3-12) in five steps in 32% overall yield from compound **90**. The main difference in overall yield being the deprotection of the TBS group.



Scheme 3-12: Benzoylation followed by DDQ deprotection of **133** providing xylose building block **132**.

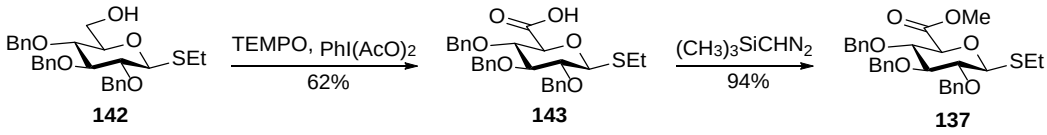
Furthermore building block **132** can be formed via the monobenzoylated by-product **94** from the regioselective benzoylation of **90** (Scheme 2-11) through levulination and PMB deprotection under the abovementioned reaction conditions without intermediate purification by flash chromatography (Scheme 3-13).



Scheme 3-13: Formation of **132** from the monobenzoylated byproduct **94** of the regioselective benzoylation.

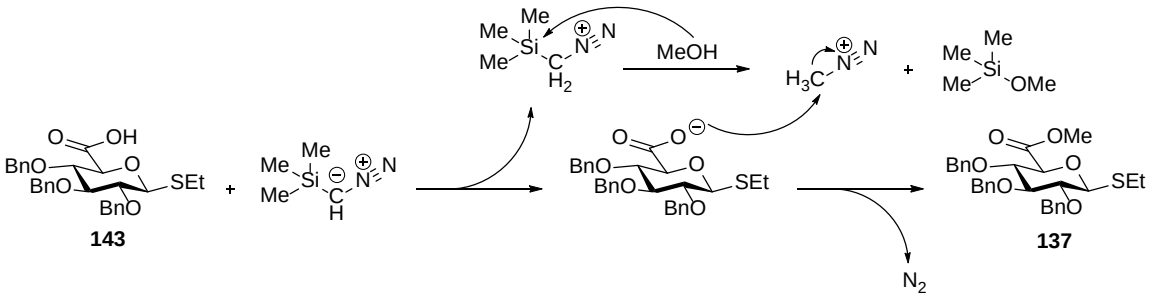
3.4 Synthesis of monomeric glucuronic acid building blocks

More material of both the two glucuronic acid building blocks **137** and **138** were produced. However, for the synthesis of **137** only the two last synthetic steps were repeated, as free 6-OH compound **142** was present in the laboratory. Consequently, **142** was subjected to oxidation using TEMPO followed by methylation using (trimethylsilyl)diazomethane according to the abovementioned conditions (Scheme 3-14).



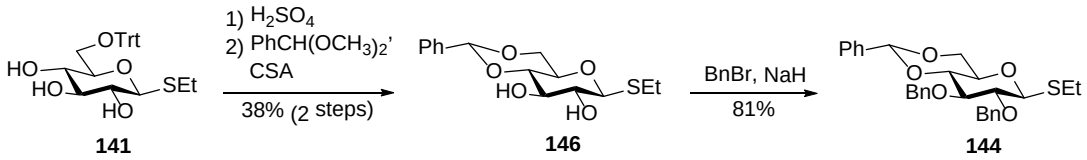
Scheme 3-14: Formation of **137** according to previously established conditions.[150]

The drop in yield observed for the TEMPO oxidation compared to the yield reported by C. d'Errico[150] cannot be explained by formation of the sulfoxide or impurities in the TEMPO reagent as also a completely new batch of the reagent was tried. Furthermore, when the same conditions were applied to the corresponding 4-OMe compound, **148**, the same drop in yield was not observed (*vide infra*). However this further proved the value of the statement by Oscarson in 2001:[76] "oxidation to an uronic acid derivative is not trivial." On the other hand, the methylation occurred with comparable result, where the reaction progressed through the proposed mechanism in Scheme 3-15. The very reactive diazomethane species can be formed *in situ* by reaction with methanol after abstraction of a hydrogen from the carboxylic acid. Subsequent attack of the carboxylate forms the methyl ester at the same time as forming nitrogen gas. The gas formation is the driving force of the reaction and at the same time makes it irreversible. It was further found, that the presence of the co-solvent toluene was essential due to poor solubility of the reactant in methanol.



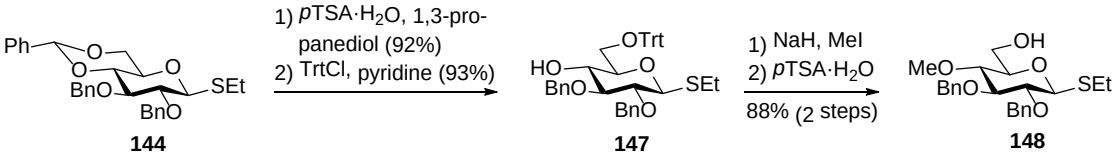
Scheme 3-15: Proposed mechanism for methylation with (trimethylsilyl)diazomethane.

As the triol **141** with an *O*-6 trityl group was present in laboratory in great amounts this compound was used as the starting material for the synthesis of glucuronic acid donor **138**. The trityl moiety was removed under acidic conditions with sulphuric acid. A benzylidene acetal between the *O*-4 and *O*-6 was then directly installed using the crude product (Scheme 3-16). Both reactions followed previously established protocols,[105,150] and although only a moderate yield is obtained for this 2-step process, the yield is comparable to the reported yield by C. d'Errico for installation of the benzylidene moiety (Scheme 3-6). Subsequent benzylation under standard conditions provided compound **144** (Scheme 3-16).



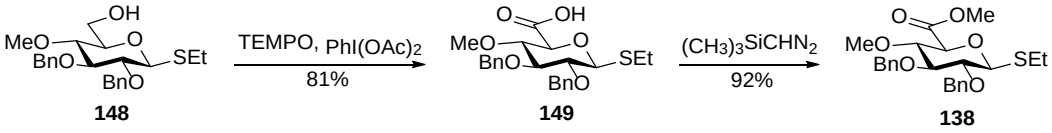
Scheme 3-16: Formation of fully protected compound **144** in three steps from compound **141**.

M. F. Böhm reported cleavage of the benzylidene acetal of compound **144** with *p*-toluenesulfonic acid (*p*TSA) monohydrate to be quick (1 h reaction time),[105] but when the deprotection was tried for this project the reaction was sluggish and full conversion was not obtained even after addition of more *p*TSA and 24 hours reaction time. As a consequence, the reaction was run in the presence of *p*TSA monohydrate and 1,3-propanediol, where the diol was added to drive the equilibrium towards the formation of a new 6-membered ring acetal. An extended reaction time was still needed, but full conversion and a good yield was now obtained (Scheme 3-17). Subsequently the trityl group was re-introduced selectively at the primary position due to the bulkiness of the substituent providing compound **147**. With compound **147** in hand the free hydroxyl group was methylated followed by hydrolysis of the trityl group in a two-step reaction as established previously (Scheme 3-17).[76,105]



Scheme 3-17: Formation of compound **148** in four steps from **144**.

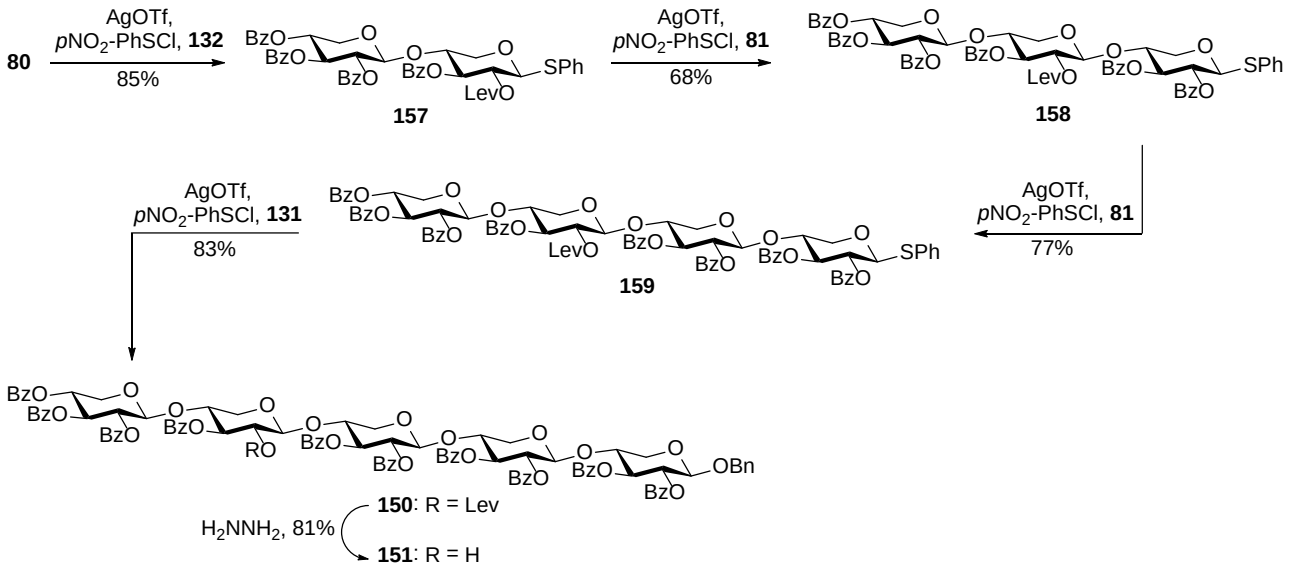
The two last synthetic transformations were the oxidation of the primary alcohol of **148** with TEMPO, followed by methylation by the same protocol as for formation of donor **137** where good yields were obtained (Scheme 3-18). It was hence possible to improve the synthesis of donor **138**, as the synthesis was completed in 45% yield over 7 steps from compound **146** compared to 30% overall yield as reported by M. F. Böhm.[105]



Scheme 3-18: The last synthetic steps for formation of glucuronic acid donor **138**.

3.5 Assembly of protected glucuronoxylan target structures

With the required building blocks in hand the common pentasaccharide backbone **150** was formed following the pre-activation glycosylation protocol (Scheme 3-19). The formation of trisaccharide **158** went in moderate yield, which is in accordance with the previous results where a Lev group is present at the 2-position of the donor (Section 2.5). Overall, it was possible to improve the yield of the four glycosylations, obtaining an overall yield of 37% (C. d'Errico 27%[150]).



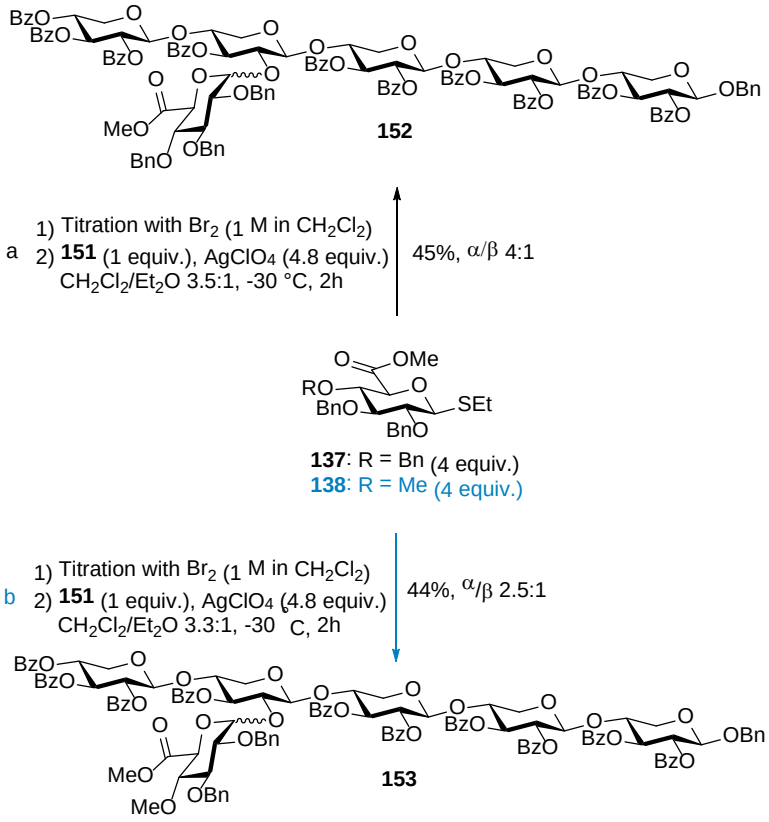
Scheme 3-19: Formation of the partially deprotected pentasaccharide backbone **151**.

Previously for synthesis of the arabinoxylans target structures (Section 2.5) it was found that the removal of multiple Lev groups worked very well and the products were obtained pure after washing. However, when the same protocol was applied for **150**, the reaction was difficult to follow on TLC due to the similarity of the retention factor of the starting material and the product. Furthermore, the reaction was sluggish and subsequent flash column chromatography was needed. The longer reaction time was probably due to low solubility of the starting material in the solvents used, even though similar solubility problems were not observed with the tetrasaccharides related to the AX fragments. Furthermore, the length and hence maybe the folding of the structure could render the Lev group less accessible and potentially also contribute to a longer reaction time. The need for flash chromatography was probably due to the formation of emulsions under the washing process, even though no explanation for the formation of these have been identified.

The β-relationship of the xylose backbone of **151** was confirmed by measurements of the $J_{C\text{-}H}$ couplings constants, which were measured to be between 162 and 165 Hz.

Unfortunately time constraints and limited amounts of material meant that the glycosylations for forming the hexasaccharides **152** and **153** could only be performed without further optimisations. As several different conditions for formation of the hexasaccharides had already been examined (Section 3.2.3), it was decided to use the conditions described in Table 3-1 Entry 11 and 12 for synthesis of **152** and **153**, respectively. Hence, **152** and **153** were synthesised in 45% and 44% yield, respectively, as inseparable mixtures (Scheme 3-20). The α/β ratio of **152**

and **153** was 4:1 and 2.5:1, respectively. Furthermore, 23% of the acceptor **151** was isolated when glycosylating with **137**, while for the corresponding glycosylation with donor **138**, acceptor **151** was reisolated in 17% yield. Hence, the yield obtained was approximately 10% lower than the previously reported yields (Table 3-1), but in both cases the α/β ratios in the glycosylations were improved.



Scheme 3-20: Formation of the hexasaccharides **152** (a) and **153** (b) according to the previously optimised glycosylation protocol.

The assignment of the major stereoisomer as the desired α-linkage was based on the chemical shift of the anomeric positions of the glucuronic acid moiety (**152**: H1: 5.33 ppm, C1: 97.3 ppm and **153**: H1: 5.25 ppm, C1: 97.6 ppm) and the measurements of $J_{C\text{-}H}$ coupling constants which for both hexasaccharides were approximately 173 Hz.

The GX target structures have been analysed by NMR making assignment of most signals in the ^{1}H and ^{13}C spectre possible by the use of 2D spectre.

3.6 Deprotection of GX targets

A few trials for deprotection of the benzoate esters in the presence of the methyl ester using Zemplén conditions had previously been run by C. d'Errico (unpublished result obtained during

a short Postdoc). It was generally found that the reaction took a long time (7 days) and more importantly, C5 at the glucuronic acid moiety had unfortunately epimerised thereby providing the β-iduronic acid. Hence, it seemed important that the methyl ester should first be deprotected separately, as the carboxylic acid should not be able to epimerise.

Several examples of deprotections of a methyl ester to form uronic acids have been reported in literature, and these all featured hydrogen peroxide and lithium hydroxide in THF with or without water in different ratios of both reagents and solvents.[166–170] In some of the examples the deprotection to the carboxylic acid was directly followed by addition of sodium hydroxide for deprotection of benzoyl or pivaloyl groups.[167,170] Based on these reports it was tried to deprotect the methyl ester of **153** using H_2O_2 (100 equiv.) and 1 M LiOH (50 equiv.) in THF. After 6 h no starting material could be seen by TLC, and the reaction was made basic (pH approximately 11) by addition of NaOH. After 24 h the mixture was neutralised, concentrated and purified using reverse phase C18 columns (Sep Pak C18 from Waters, H_2O + 40% acetonitrile was needed for elution). Due to time limitations and low amounts obtained full NMR characterisation was not performed. However, limited NMR analysis was completed for the two major fractions isolated. Generally the analysis showed that the compound in the fractions only contained one methyl group (δ_H 3.53 ppm, δ_C 60 ppm), as well as the carbonyl signal had moved from δ_C 170.3 ppm for the methyl ester in **153** to approximately δ_C 174 ppm for the new compounds, which thereby could represent the carbonyl of a carboxylic acid. In HMBC it could be seen that a peak at δ_H 4.66 ppm (d, J = 9.8 Hz) and δ_C 70.7 ppm coupled to the carbonyl (δ_C 174 ppm), which would correspond to the 5-position in the glucuronic acid. Both HSQC and HMBC indicated the presence and the chemical shift of the 5-position, but for one of the fractions the peak (δ_H 4.66 ppm) could not be observed in ^{1}H NMR, for which no explanation has been found as of now. However, importantly the α-product seemed to be the major product, as a crosspeak in HSQC at δ_H 5.46 ppm (J = 3.4 Hz) and δ_C 96.4 ppm could be seen, corresponding well to the 1-position in a glucuronic acid. Furthermore, MS (MALDI) of both fractions confirmed the mass of the product (major peaks M+Na^+ and M+2Na^+-H (sodium salt of the carboxylic acid)).

Together these data indicated that the correct product had been obtained in good yield (91% collectively) over 2 steps. Furthermore, the data do not indicate epimisation to the β-iduronic acid, as a large coupling constant is observed for what is believed to be H5. Additionally the

chemical shift of what is believed to be C1 of the glucuronic acid was 96.4 ppm and not approximately 100 ppm which would be expected for the epimerised product.[171–173]

3.7 Future work

The time limitations of this project meant that no further trials with the deprotections of the obtained hexasaccharides **152** and **153** were accomplished at the deadline of this thesis. But, it is the hope that the syntheses of these glucuronoxylan fragments will be finalised shortly.

The two-step one-pot deprotection of the methyl ester followed by removal of the benzoyl groups should be tried again using **153** as the starting material, where the product should be fully characterised to ascertain the identity. If the product identity can be confirmed, then it should be possible to use the same conditions for the deprotections of **152**. For the trials the partly deprotected product would need purifications, although, it would be preferable to subject the partly deprotected crude product directly to the following hydrogenolysis for removal of the benzyl groups, and then only perform one purification after the final step.

For the hydrogenolysis several different palladium sources could be tried, but based on published hydrogenloysis of compounds containing glucuronic acids by Oscarson and co-workers[76,104] the first choice would be Pearlman's catalyst in a mixture of an alcohol (MeOH or EtOH) and acetic acid. For the final purification PD MiniTrap G-10 would be a possibility also based on the final purification utilised by Oscarson and co-workers.[76,104]

3.8 Concluding remarks

Two glucuronoxylan fragments had been envisioned as the targets for this project already before the start of the work for this thesis. However, due to lack of materials most of the previous synthetic work was repeated. The synthetic strategy for two glucuronoxylan fragments were a further development of the strategy developed for the arabinoxylan fragments. But the synthesis of another xylopyranose building block **132** was needed to be able to install the glucuronic acid in the 2-position of the xylose unit. This xylose building block, **132**, was formed in five steps from the common intermediate, i.e. the 4-*O*-PMB protected thioxylopyranose **90**, as also previously utilised for the syntheses of the other xylopyranose building blocks. Furthermore, the syntheses of two glucuronic acid building blocks were completed, where the building blocks **137** and **138** were derived from thioglucopyranose **140** in five and eight steps, respectively. For this thesis many of the synthetic steps for the syntheses of **137** and **138** were reproduced, and in some reactions with improved yields.

The xylose pentasaccharide backbone was synthesised following the linear, iterative strategy in an improved overall yield (37%) compared to previous reports from the group (27% overall yield). After deprotection of the Lev group the hexasaccharides **152** and **153** were formed in 45% (α/β 4:1) and 44% (α/β 2.5:1) yield, respectively. The yields obtained for these glycosylations were lower than reported in the previous trial in the group, but the α/β ratio was improved in both cases.

Promising results were obtained for the selective removal of the methyl ester to afford the free carboxylic acid of the glucuronic acid followed by removal of the benzoyl groups under basic conditions. Further trials for these reactions are needed, as well as the remaining benzyl groups should be removed through hydrogenolysis.

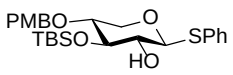
Possibly some of the problems with deprotection, especially the epimerisation observed during direct benzoyl deprotection, could have been avoided using a post-glycosylation oxidation approach for the glucuronic acid. However, with the latest results the syntheses of the two glucuronoxylan fragment are expected to be finished in due course.

3.9 Experimental

General considerations

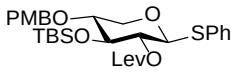
Starting materials, reagents, and solvents were purchased from commercial suppliers and were used without further purification unless otherwise noted. All solvents used were of analytical HPLC grade. Anhydrous solvents were obtained from an Innovative Technology PS-MD-7 PureSolv solvent purification system. Except for dichloromethane and toluene for glycosylations which were dried over 3 Å molecular sieves prior to use.[144] Air- and moisture sensitive experiments were conducted under inert atmosphere (N_2) in dried glassware and anhydrous solvents. Water content was checked by 899 Coulometer from Metrohm. Crushed 3 Å molecular sieves for glycosylations were stored in the oven at 110 °C, before activation with heating under vacuum. Evaporation of solvents was performed with a VWR International Laborota 400 under reduced pressure (*in vacuo*) at temperatures between 23-50 °C. Thin-layer chromatography (TLC) was carried out using Merck Aluminium Sheets pre-coated with 0.25 mm silica gel, C-60 F_{254} plates. TLC plates were inspected under UV light or visualized by charring after dipping in a cerium ammoniumsulfate solution (1% cerium(IV)sulphate and 2.5% ammonium heptamolybdate in a 10% sulfuric acid solution). Flash column chromatography was performed using Merck Geduran silica gel 60 Å (40-63 μm) as the stationary phase. Eluent systems are specified for each R_f-value, and ratios are given as volume ratios. NMR spectre were recorded with a Bruker Ascend™ 400 (400 MHz and 101 MHz) with a Prodigy cyroprobe. The recorded chemical shift are reported in parts per million (δ = scale) relative to the residual solvent peak in $CDCl_3$ (δ_H = 7.26 ppm, δ_C = 77.16 ppm), methanol-*d4* (δ_H = 4.87 ppm, δ_C = 49.00 ppm), and D_2O (δ_H = 4.79 ppm). The annotation (2C), (3C) ect. indicates when multiple carbon signals are overlapping. Coupling constants (*J*) are reported in Hz. Multiplicities are reported as singlet (s), broad singlet (bs), doublet (d), doublets of doublets (dd), doublets of doublets of doublets (ddd), triplet (t), triplets of doublets (td), quartet (q), and multiplet (m). Assignment of ^1H and ^{13}C resonances were based on APT, DQF-COSY, HSQC, H2BC, HMBC, TOCSY, and HSQC-TOSCY when needed. Optical rotation was measured on a Perkin Elmer Model 341 Polarimeter. HRMS analysis was performed on either a UHPLC-QTOF system (Dionex ultimate 3000 and Bruker MaXis) with an electrospray ionization (ESI) source or a MALDI-TOF system (Bruker Solarix XR 7T) and controlled using Data Analysis 4.2 software. New compounds have been characterised by R_f, $[\alpha]_D$, ^1H NMR, ^{13}C NMR, and HRMS.

Phenyl 3-*O*-*tert*-butyldimethylsilyl-4-*O*-(*p*-methoxy)benzyl-1-thio-β-D-xylopyranoside (135)



Diol **90** (2.073 g, 5.72 mmol, 1.0 equiv.) was dissolved in dry THF (63.0 mL) and cooled to -78 °C. 2,6-Lutidine (1.00 mL, 8.85 mmol, 1.5 equiv.) and TBSOTf (1.60 mL, 6.96 mmol, 1.2 equiv.) were added. Stirring was continued for 20 min until TLC (heptane/EtOAc 1:1) showed full conversion. The reaction mixture was diluted with CH_2Cl_2 and washed with H_2O, dried over Na_2SO_4, filtered, and conc. *in vacuo*. The crude product was purified by flash column chromatography (heptane/EtOAc 10:1). The product was obtained as a clear oil (2.220 g, 81%). $\mathbf{R_f}$(heptane/EtOAc 9:1) 0.28. $[\boldsymbol{\alpha}]_{\boldsymbol{D}}^{\boldsymbol{298\,K}}$ = -130.2 (*c* 0.083, $CHCl_3$). **^{1}H NMR** (400 MHz, $CDCl_3$) δ 7.49 – 7.47 (m, 1H, Ar*H*), 7.30 – 7.20 (m, 7H, Ar*H*), 6.89 – 6.87 (m, 1H, Ar*H*), 5.20 (d, $J_{1,2}$ = 3.0 Hz, 1H, H1), 4.62 (d, J_{gem} = 11.5 Hz, 1H, OCH_2Ph), 4.53 (d, J_{gem} = 11.5 Hz, 1H, OCH_2Ph), 4.40 (dd, $J_{5,5'}$ = 12.4 Hz, $J_{5,4}$ = 1.8 Hz, 1H, H5), 3.87 (t, *J* = 4.4 Hz, 1H, H3), 3.81 (s, 3H, OCH_3), 3.75 (d, *J* = 9.4 Hz, 1H, OH), 3.70 – 3.66 (m, 1H, H2), 3.62 (dd, $J_{5',5}$ = 12.4 Hz, $J_{5',4}$ = 3.7 Hz, 1H, H5'), 3.34 (q, *J* = 5.4 Hz, 1H, H4), 0.93 (s, 9H, $C(CH_3)_3$), 0.12 (s, 3H, $SiCH_3$), 0.05 (s, 3H, $SiCH_3$). **^{13}C NMR** (101 MHz, $CDCl_3$) δ 159.5 (1C, Ar*C*), 136.6 (1C, Ar*C*), 130.8 (2C, Ar*C*H), 129.5 (2C, Ar*C*H), 129.4 (1C, Ar*C*), 128.8 (2C, Ar*C*H), 126.9 (1c, Ar*C*H), 114.0 (2C, Ar*C*H), 88.9 (C1), 76.0 (C4), 72.5 (C2), 71.4 (OCH_2Ph), 69.5 (C3), 59.9 (C5), 55.3 (OCH_3), 25.8 ($C(CH_3)_3$), 18.1 ($C(CH_3)_3$), -4.8 ($SiCH_3$), -5.0 ($SiCH_3$). **HRMS:** calcd. for $C_{25}H_{36}O_5SSiNa^+$ *m/z* 499.1944; found *m/z* 499.1956.

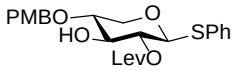
Phenyl 3-*O*-*tert*-butyldimethylsilyl-2-*O*-levulinoyl-4-*O*-(*p*-methoxy)benzyl-1-thio-β-D-xylopyranoside (136)



Alcohol **135** (0.862 g, 1.81 mmol, 1.0 equiv.) was dissolved in CH_2Cl_2 (36.0 mL) followed by addition of DCC (1.119 g, 5.43 mmol, 3.0 equiv.), DMAP (0.332 g, 2.72 mmol, 1.5 equiv.), and LevOH (0.315 g, 2.72 mmol, 1.5 equiv.) and the mixture was left to stir at r.t. overnight. The formation of a white precipitate was observed, although full conversion was not obtained as observed by TLC (heptane/EtOAc 1:1). The reaction mixture was filtered through a pad of Celite, the filtrate was conc. *in vacuo*, and the crude product purified by flash column chromatography (heptane/EtOAc 7:1). The product was obtained as a colourless oil (0.828 g, 80 %) alongside re-isolated starting material **135** (0.108 g, 12%). $\mathbf{R_f}$(heptane/EtOAc 3:2) 0.27. $[\boldsymbol{\alpha}]_{\boldsymbol{D}}^{\boldsymbol{298\,K}}$ = -33.2 (*c* 0.28, $CHCl_3$). **^{1}H NMR** (400 MHz, $CDCl_3$) δ 7.46

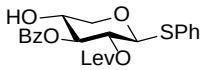
– 7.44 (m, 2H, Ar*H*), 7.30 – 7.23 (m, 5H, Ar*H*), 6.87 – 6.85 (m, 2H, Ar*H*), 4.89 – 4.85 (m, 2H, H1, H2), 4.58 (d, J_{gem} = 11.6 Hz, 1H, OCH_2Ph), 4.47 (d, J_{gem} = 11.6 Hz, 1H, OCH_2PH), 4.13 (q, J = 7.0 Hz, 1H, H5), 3.80 (s, 3H, OCH_3), 3.80 – 3.77 (m, 1H, H3), 3.39 – 3.33 (m, 2H, H4, H5'), 2.78 – 2.74 (m, 2H, COCH_2CH_2), 2.66 – 2.61 (m, 2H, CH_2CH_2COO), 2.17 (s, 3H, CH_3CO), 0.89 (s, 9H, C(CH_3)$_3$), 0.09 (s, 3H, SiCH_3), 0.06 (s, 3H, SiCH_3). **^{13}C NMR** (101 MHz, $CDCl_3$) δ 206.2 (1C, C_q, CO, Lev), 171.7 (1C, C_q, COO, Lev), 159.3 (1C, Ar*C*), 135.0 (1C, Ar*C*), 131.2 (2C, Ar*C*H), 130.1 (1C, Ar*C*), 129.5 (2C, Ar*C*H), 128.9 (2C, Ar*C*H), 127.3 (1C, Ar*C*H), 113.8 (2C, Ar*C*H), 86.7 (C1), 76.6 (C4), 72.8 (C2), 72.9 (O*C*H$_2$Ph), 71.8 (C3), 64.4 (C5), 55.3 (*C*H$_3$CO), 37.9 (CO*C*H$_2$CH$_2$), 29.9 (*C*H$_3$CO), 28.3 (CH$_2$*C*H$_2$COO), 25.7 (C(*C*H$_3$)$_3$), 18.0 (*C*(CH$_3$)$_3$), -4.5 (Si*C*H$_3$), -4.8 (Si*C*H$_3$). **HRMS:** calcd. for $C_{30}H_{42}O_7SSiNa^+$ *m/z* 597.2312; found *m/z* 597.2325.

Phenyl 2-*O*-levulinoyl-4-*O*-(*p*-methoxy)benzyl-1-thio-β-D-xylopyranoside (133)



A mixture of **136** (0.177 g, 0.31 mmol, 1.0 equiv.) in CH_3CN (90 mL) was cooled to 0 °C. A 20% HF solution (0.33 mL, 6.17 mmol, 20.0 equiv.) was added and the reaction mixture was allowed to attain r.t. overnight. TLC (heptane/EtOAc 1:1) showed full conversion, and afterwards remaining HF was quenched by addition of methoxytrimethylsilane (1.70 mL, 12.32 mmol, 40 equiv.) and stirred for 1 h. The mixture was diluted with CH_2Cl_2, washed with sat. aq. $NaHCO_3$ and water, dried over Na_2SO_4, filtered, and the solvent removed *in vacuo*. The crude mixture was purified by flash column chromatography (heptane/EtOAc 4:3) yielding the product **133** as a white amorphous solid (0.087 g, 61%). $\mathbf{R_f}$(heptane/EtOAc 1:1) 0.24. $[\boldsymbol{\alpha}]_{\boldsymbol{D}}^{\boldsymbol{298\,K}}$ = -18.8 (*c* 0.47, $CHCl_3$). **^{1}H NMR** (400 MHz, $CDCl_3$) δ 7.47 – 7.45 (m, 2H, Ar*H*), 7.32 – 7.24 (m, 5H, Ar*H*), 6.89 – 6.83 (m, 2H, Ar*H*), 4.83 (dd, $J_{2,1}$ = 9.8 Hz, $J_{2,3}$ = 9.0 Hz, 1H, H2), 4.69 (d, J_{gem} = 11.5 Hz, 1H, OCH_2Ph), 4.60 (d, $J_{1,2}$ = 9.8 Hz, 1H, H1), 4.58 (d, J_{gem} = 11.4 Hz, 1H, OCH_2Ph), 4.01 (dd, $J_{5,5'}$ = 11.5 Hz, $J_{5,4}$ = 5.2 Hz, 1H, H5$_{eq}$), 3.80 (s, 3H, OCH_3), 3.75 (t, J = 8.9 Hz, 1H, H3), 3.53 (ddd, $J_{4,5'}$ = 10.2 Hz, $J_{4,3}$ = 8.9 Hz, $J_{4,5}$ = 5.2 Hz, 1H, H4), 3.22 (dd, $J_{5',5}$ = 11.5 Hz, $J_{5',4}$ = 10.3 Hz, 1H, H5'$_{ax}$), 2.94 – 2.76 (m, 2H, COCH_2CH$_2$), 2.69 – 2.56 (m, 2H, CH$_2$CH_2COO), 2.20 (s, 3H, CH_3CO). **^{13}C NMR** (101 MHz, $CDCl_3$) δ 207.6 (1C, C_q, CO, Lev), 172.1 (1C, C_q, COO, Lev), 159.6 (1C, Ar*C*), 132.9 (1C, Ar*C*), 132.5 (2C, Ar*C*H), 130.2 (1C, Ar*C*), 129.7 (2C, Ar*C*H), 129.1 (2C, Ar*C*H), 128.0 (1C, Ar*C*H), 114.1 (2C, Ar*C*H), 86.5 (C1), 76.6 (C4), 76.4 (C3), 73.2 (O*C*H$_2$Ph), 72.8 (C2), 67.9 (C5), 55.4 (O*C*H$_3$), 38.5 (COCH_2*C*H$_2$), 30.0 (*C*H$_3$CO), 28.4 (CH$_2$*C*H$_2$COO). **HRMS:** calcd. for $C_{24}H_{28}O_7SNa^+$ *m/z* 483.1447; found *m/z* 483.1458.

Phenyl 3-*O*-benzoyl-2-*O*-levulinoyl-1-thio-β-D-xylopyranoside (132)

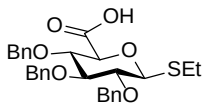


To alcohol **133** (0.893 g, 1.94 mmol, 1.0 equiv.) dissolved in dry pyridine (4.0 mL) was added BzCl (0.34 mL, 2.91 mmol, 1.5 equiv.) and the solution was left to stir at r.t. After 17 h TLC (heptane/EtOAc 1:1) showed full conversion. DMAPA (0.37 mL, 2.91 mmol, 1.5 equiv.) was added to quench left over BzCl. After stirring for 1 h CH_2Cl_2 was added and the mixture was washed with 1 M HCl and brine, dried over $MgSO_4$, filtered, and conc. *in vacuo* providing the crude as a yellow oil (1.0053 g, 1.78 mmol). The crude product was directly dissolved in a 10:1 mixture of CH_2Cl_2/H_2O (11 mL), DDQ (0.566 g, 2.49 mmol, 1.4 equiv.) was added and the mixture stirred vigorously at r.t. in the dark overnight. TLC (heptane/EtOAc 1:1) showed full conversion, and the reaction mixture was filtered through a thick pad of Celite. The filtrate was washed with sat. aq. $NaHCO_3$, and brine, dried over $MgSO_4$, filtered, and conc. *in vacuo*. The residue was purified by flash column chromatography (hexane/EtOAc 1:1) to obtain the product **131** as a white solid (0.696 g, 81 %). $\mathbf{R_f}$(hexane/EtOAc 2:1) 0.42. $[\boldsymbol{\alpha}]_{\boldsymbol{D}}^{\boldsymbol{298\,K}}$ = +16.1 (*c* 0.6, $CHCl_3$). **^{1}H NMR** (400 MHz, $CDCl_3$) δ 8.08 – 8.05 (m, 2H, Ar*H*), 7.62 – 7.58 (m, 1H, Ar$H_{para(Bz)}$), 7.53 – 7.51 (m, 2H, Ar*H*), 7.48 – 7.45 (m, 2H, Ar*H*), 7.36 – 7.31 (m, 3H, Ar*H*), 5.19 – 5.15 (m, 2H, H2, H3), 4.96 – 4.91 (m, 1H, H1), 4.37 (dd, $J_{5eq,5ax}$ = 11.9 Hz, $J_{5eq,4}$ = 4.4 Hz, 1H, $H5_{eq}$), 3.93 – 3.88 (m, 1H, H4), 3.54 (dd, $J_{5ax,5eq}$ = 11.9 Hz, $J_{5ax,4}$ = 7.9 Hz, 1H, $H5_{ax}$), 3.03 (bs, 1H, OH), 2.71 – 2.67 (m, 2H, COCH_2CH$_2$), 2.64 – 2.50 (m, 2H, CH$_2$CH_2COO), 2.09 (s, 3H, CH_3CO). **^{13}C NMR** (101 MHz, $CDCl_3$) δ 206.0 (1C, C_q, CO, Lev), 171.4 (1C, C_q, COO, Lev), 167.0 (1C, C_q, CO, Bz), 133.9 (1C, ArC_{para}H$_{(Bz)}$), 133.0 (1C, Ar$C_{ipso(SPh)}$), 132.7 (2C, Ar*C*H), 130.3 (2C, Ar*C*H), 129.2 (2C, Ar*C*H), 129.0 (1C, Ar$C_{ipso(Bz)}$), 128.7 (2C, Ar*C*H), 128.2 (1C, ArC_{para}H$_{(SPh)}$H), 86.5 (C1), 75.8, 69.8 (C2,C3), 68.3 (C4), 67.5 (C5), 38.0 (CO*C*H$_2$CH$_2$), 29.8 (*C*H$_3$CO), 28.1 (CH$_2$*C*H$_2$COO). **HRMS:** calcd. for $C_{23}H_{24}O_7SNa^+$ *m/z* 467.1134; found *m/z* 467.1146.

Phenyl 3-*O*-benzoyl-2-*O*-levulinoyl-1-thio-β-D-xylopyranoside (132): Alcohol **94** (0.286 g, 0.64 mmol, 1.0 equiv.) and 4-oxopentanoic acid (0.084 g, 0.72 mmol, 1.1 equiv.) in CH_2Cl_2 (1.8 mL) were cooled to 0 °C and a solution of DMAP (0.010 g, 0.081 mmol, 0.13 equiv.) and DDC (1.190 g, 0.92 mmol, 1.4 equiv.) in CH_2Cl_2 (0.4 mL) was added. The solution was left to stir at r.t. overnight. TLC analysis (hexane/EtOAc 4:1) showed the consumption of the starting material and formation of product. The reaction mixture was filtered and conc. *in vacuo.* The residue was dissolved in CH_2Cl_2/H_2O 10.1 (10.5 mL) and stirred vigorously. DDQ (0.204 g, 0.90 mmol, 1.4 equiv.) was added and the mixture was left to stir at r.t. overnight. TLC analysis

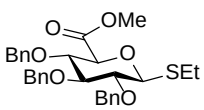
(hexane/EtOAc 2:1) showed the consumption of the starting material and formation of product. The solution was filtered through a big pad of Celite and washed with aq. sat. $NaHCO_3$ and brine, dried over Na_2SO_4, filtered, and conc. *in vacuo*. The residue was purified by flash column chromatography (hexane/EtOAc 3:1) to provide the product as a white crystalline solid (0.197 g, 72% over 2 steps).

Ethyl 2,3,4-tri-*O*-benzyl-1-thio-β-D-glucopyranosiduronic acid (143)



Primary alcohol **142** (1.997 g, 4.04 mmol, 1.0 equiv.) was dissolved in a solvent mixture of CH_2Cl_2 and H_2O 2:1 (20 mL) and stirred vigorously. TEMPO (0.122 g, 0.81 mmol, 0.2 equiv.) and $PhI(OAc)_2$ (3.253 g, 10.10 mmol, 2.5 equiv.) were added. Full conversion of the starting material was seen after 60 min with TLC (heptane/EtOAc 1:1). The remaining oxidant was quenched by addition of 10% aq. $Na_2S_2O_3$ (80 mL). The water phase was extracted with EtOAc, and the organic layer was dried over Na_2SO_4, filtered, and conc. *in vacuo*. The carboxylic acid **143** (1.275 g, 62 %) was obtained by purification by flash column chromatography (heptane/EtOAc/AcOH 7:1:0.25). $\mathbf{R_f}$(heptane/EtOAc/AcOH 7:3:0.5) 0.40. $[\boldsymbol{\alpha}]_{\boldsymbol{D}}^{\boldsymbol{298\ K}}$ = -10.5 (*c* 1, $CHCl_3$). **^{1}H NMR** (400 MHz, $CDCl_3$) δ 7.37 – 7.22 (m, 15H), 4.90 (d, *J* = 10.3 Hz, 1H, C2-OC*H*$_2$Ph), 4.86 (d, *J* = 11.1 Hz, 1H, C3-OC*H*$_2$Ph), 4.81 (d, *J* = 11.0 Hz, 1H, C3-OC*H*$_2$Ph), 4.77 (d, *J* = 10.7 Hz, 1H, C4-OC*H*$_2$Ph), 4.72 (d, *J* = 10.3 Hz, 1H, C2-OC*H*$_2$Ph), 4.65 (d, *J* = 10.7 Hz, 1H, C4-OC*H*$_2$Ph), 4.58 (d, $J_{1,2}$ = 9.6 Hz, 1H, H1), 3.97 (d, $J_{5,4}$ = 9.2 Hz, 1H, H5), 3.83 (t, *J* = 8.9 Hz, 1H, H4), 3.71 (t, *J* = 8.5 Hz, 1H, H3), 3.48 (dd, $J_{2,1}$ = 9.6 Hz, $J_{2,3}$ = 8.3 Hz, 1H, H2), 2.83 – 2.68 (m, 2H, SC*H*$_2$CH$_3$), 1.31 (t, *J* = 7.4 Hz, 3H, SCH$_2$C*H*$_3$). **^{13}C NMR** (101 MHz, $CDCl_3$) δ 171.9 (1C, C_q, *C*OOH), 138.2 (1C, Ar*C*), 137.8 (1C, Ar*C*), 137.4 (1C, Ar*C*), 128.6 (2C, Ar*C*H), 128.6 (2C, Ar*C*H), 128.6 (2C, Ar*C*H), 128.4 (2C, Ar*C*H), 128.3 (2C, Ar*C*H), 128.2 (1C, Ar*C*H), 128.1 (1C, Ar*C*H), 128.0 (1C, Ar*C*H), 127.9 (2C, Ar*C*H), 85.7 (C1), 85.2 (C3), 81.1 (C2), 78.8 (C4), 77.3 (C5), 75.8 (C3-O*C*H$_2$Ph), 75.6 (C2-O*C*H$_2$Ph), 75.2 (C4-O*C*H$_2$Ph), 25.5 (S*C*H$_2$CH$_3$), 15.2 (SCH$_2$*C*H$_3$). NMR data are in accordance with literature values.[174]

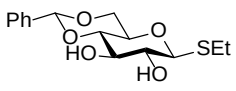
Methyl (ethyl 2,3,4-tri-*O*-benzyl-1-thio-β-D-glucopyranoside) uronate (137)



Carboxylic acid **143** (0.152 g, 0.30 mmol, 1.0 equiv.) was dissolved in a solvent mixture of dry MeOH and toluene 1:1 (4.5 mL) and $(CH_3)_3SiCHN_2$ 2 M in hexane

(0.59 mL, 1.18 mmol, 4.0 equiv.) was slowly added. After 3 h TLC (heptane/EtOAc 1:1) showed full conversion of the starting material, and the excess reagent was quenched by addition of AcOH (1.5 mL). The mixture was conc. *in vacuo* and the crude material purified by flash column chromatography (heptane/EtOAc 4:1) affording the methyl ester **137** as a white amorphous solid (0.145 g, 94%). $\mathbf{R_f}$(heptane/EtOAc 4:1) 0.39. $[\boldsymbol{\alpha}]_{\boldsymbol{D}}^{\boldsymbol{298\ K}}$ = -7.6 (*c* 1, $CHCl_3$). **^{1}H NMR** (400 MHz, $CDCl_3$) δ 7.38 – 7.27 (m, 13H, Ar*H*), 7.24 – 7.21 (m, 2H, Ar*H*), 4.93 – 4.90 (m, 2H, OCH_2Ph), 4.85 (d, *J* = 11.0 Hz, 1H, C3-OCH_2Ph), 4.79 (d, *J* = 10.8 Hz, 1H, C4-OCH_2Ph), 4.74 (d, *J* = 10.2 Hz, 1H, C2-OCH_2Ph), 4.61 (d, *J* = 10.8 Hz, 1H, C4-OCH_2Ph), 4.51 (d, $J_{1,2}$ = 9.7 Hz, 1H, H1), 3.90 (d, $J_{5,4}$ = 9.7 Hz, 1H, H5), 3.87 – 3.82 (m, 1H, H4), 3.72 – 3.68 (m, 4H, H3, COOCH_3), 3.48 (dd, $J_{2,1}$ = 9.6 Hz, $J_{2,3}$ = 8.9 Hz, 1H, H2), 2.83 – 2.68 (m, 2H, SC$H_2$$CH_3$), 1.31 (t, *J* = 7.4 Hz, 3H, SCH_2CH_3). **^{13}C NMR** (101 MHz, $CDCl_3$) δ 168.9 (1C, C_q, COO), 138.4 (1C, Ar*C*), 137.9 (1C, Ar*C*), 137.9 (1C, Ar*C*), 128.6 (2C, Ar*C*H), 128.6 (4C, Ar*C*H), 128.4 (2C, Ar*C*H), 128.1 (2C, Ar*C*H), 128.1 (1C, Ar*C*H), 128.0 (1C, Ar*C*H), 127.9 (2C, Ar*C*H), 127.9 (1C, Ar*C*H), 86.0 (C1/C3), 86.0 (C1/C3), 81.3 (C2), 79.4 (C4), 78.2 (C5), 76.0 (C4-O*C*H_2Ph), 75.7 (C3-O*C*H_2Ph), 75.2 (C2-O*C*H_2Ph), 52.6 (COO*C*H_3), 25.3 (S*C*$H_2$$CH_3$), 15.1 ($SCH_2$*C*$H_3$). **HRMS:** calcd. for $C_{30}H_{34}O_6SH^+$ *m/z* 523.2149; found *m/z* 523.2157. NMR data are in accordance with literature values.[174]

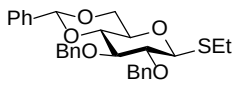
Ethyl 4,6-*O*-benzylidene-1-thio-β-D-glucopyranoside (146)



To triol **141** (8.000 g, 17.15 mmol, 1.0 equiv.) dissolved in MeOH (25.0 mL) was added concentrated H_2SO_4 (0.18 mL, 3.43 mmol, 0.2 equiv.). Full conversion of the starting material was observed by TLC (EtOAc) after stirring at r.t. for 45 min. The acid was quenched by addition of Na_2CO_3 (1.280 g, 12.08 mmol, 0.7 equiv.) by stirring for 2 h. Amberlite H^+ was added to removed left over base and after stirring for 1.5 h the mixture was filtered through a pad of Celite and conc. *in vacuo*. The crude residue was used directly for the next reaction. The crude was dissolved in dry CH_3CN (120.0 mL) and $PhCH(OMe)_2$ (8.0 mL, 53.3 mmol, 3.1 equiv.) and CSA (2.005 g, 8.63 mmol, 0.5 equiv.) were added. The mixture was stirred at r.t. for 7.5 h at which point TLC (EtOAc/acetone 2:1) showed full conversion of the starting material. The mixture was neutralized by addition of Et_3N (1.7 mL, 12.19 mmol, 0.7 equiv.), filtered and the solvent evaporated *in vacuo*. The crude was purified by flash column chromatography (heptane/EtOAc 6:4) to yield the product as a white amorphous solid (2.016 g, 38%). $\mathbf{R_f}$(heptane/EtOAc 5:1) 0.54. **^{1}H NMR** (400 MHz, $CDCl_3$) δ 7.50 – 7.48 (m, 2H, Ar*H*), 7.39 – 7.36 (m, 3H, Ar*H*), 5.53 (s, 1H, C*H*Ph), 4.45 (d, *J* = 9.8 Hz, 1H, H1), 4.34 (dd, *J* = 10.5, 4.8 Hz, 1H,

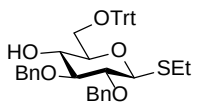
H6), 3.81 (t, *J* = 8.8 Hz, 1H, H3), 3.76 (t, *J* = 10.1 Hz, 1H, H6'), 3.56 (t, *J* = 9.2 Hz, 1H, H4), 3.53 – 3.43 (m, 2H, H2, H5), 2.75 (qd, *J* = 7.4, 2.0, 2H, SCH_2CH$_3$), 1.32 (t, *J* = 7.4 Hz, 3H, SCH$_2$CH_3). **^{13}C NMR** (101 MHz, $CDCl_3$) δ 137.0 (1C, ArC_{ipso}), 129.5 (1C, ArC_{para}H), 128.5 (2C, ArC_{meta}H), 126.4 (2C, ArC_{ortho}H), 102.1 (*C*HPh), 86.7 (C1), 80.5 (C4), 74.7 (C3), 73.3 (C2), 70.7 (C5), 68.7 (C6), 24.9 (S*C*H$_2$CH$_3$), 15.4 (SCH$_2$*C*H$_3$). NMR data are in accordance with literature values.[175]

Ethyl 2,3-di-*O*-benzyl-4,6-*O*-benzylidene-1-thio-β-D-glucopyranoside (144)



Diol **146** (2.021 g, 6.44 mmol, 1.0 equiv.) was dissolved in dry DMF (58.0 mL) and cooled to 0 °C. At this point NaH 60% in oil dispersion (0.983 g, 25.66 mmol, 4.0 equiv.) was added. The mixture was stirred for 10 min. BnBr (1.90 mL, 16.10 mmol, 2.5 equiv.) was added and the temperature was raised to r.t. The reaction mixture was stirred for 17 h, and the remaining NaH and BnBr were quenched by addition of MeOH (15 mL). The mixture was diluted with Et_2O and the organic phase was washed with H_2O, dried over Na_2SO_4, filtered, conc. *in vacuo*, and purified by flash column chromatography (heptane/EtOAc 9:1) to afford the product as a white solid (2.586 g, 81%). $\mathbf{R_f}$(heptane/EtOAc 3:2) 0.57. $[\boldsymbol{\alpha}]_{\boldsymbol{D}}^{\boldsymbol{298\ K}}$ = -41.5 (*c* 1, $CHCl_3$). **^{1}H NMR** (400 MHz, $CDCl_3$) δ 7.51 – 7.49 (m, 2H, Ar*H*), 7.42 – 7.28 (m, 13H, Ar*H*), 5.59 (s, 1H, C*H*Ph), 4.96 (d, *J* = 11.3 Hz, 1H, OCH_2Ph), 4.90 (d, *J* = 10.2 Hz, 1H, OCH_2Ph), 4.84 – 4.80 (m, 2H, OCH_2Ph), 4.58 (d, *J* = 9.8 Hz, 1H, H1), 4.37 (dd, *J* = 10.5, 5.0 Hz, 1H, H6), 3.87 – 3.74 (m, 2H, H3, H6), 3.73 (t, *J* = 9.3 Hz, 1H, H4), 3.50 – 3.43 (m, 2H, H2, H5), 2.84 – 2.70 (m, 2H, SCH_2CH$_3$), 1.33 (t, *J* = 7.4 Hz, 3H, SCH$_2$CH_3). **^{13}C NMR** (101 MHz, $CDCl_3$) δ 138.5 (1C, Ar*C*), 138.1 (1C, Ar*C*), 137.4 (1C Ar*C*), 129.1 (1C, Ar*C*H), 128.5 (4C, Ar*C*H), 128.5 (2C, Ar*C*H), 128.4 (2C, Ar*C*H), 128.2 (2C, Ar*C*H), 128.0 (1C, Ar*C*H), 127.8 (1C, Ar*C*H), 126.1 (2C, Ar*C*H), 101.3 (*C*HPh), 86.0 (C1), 82.9 (C3), 81.7 (C4), 81.4 (C2), 76.1 (O*C*H$_2$Ph), 75.4 (O*C*H$_2$Ph), 70.4 (C5), 68.9 (C6), 25.3 (S*C*H$_2$CH$_3$), 15.3 (SCH$_2$*C*H$_3$). **HRMS:** calcd. for $C_{29}H_{32}O_5SH^+$ *m/z* 493.2043; found *m/z* 493.2050. Data are in accordance with literature values.[176,177]

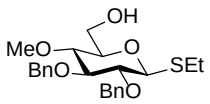
Ethyl 2,3-di-*O*-benzyl-6-*O*-trityl-1-thio-β-D-glucopyranoside (147)



Acetal **144** (2.333 g, 4.74 mmol, 1.0 equiv.) was dissolved in a 1:1 mixture of CH_2Cl_2 and MeOH (12.0 mL) and *p*TSA·H_2O (0.272 g, 1.43 mmol, 0.3 equiv.) and 1,3-propanediol (1.70 mL, 23.5 mmol, 5.0 equiv.) were added, followed by

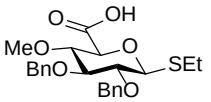
stirring overnight at r.t. TLC analysis (heptane/EtOAc 3:2) showed full conversion. The mixture was neutralized by addition of sat. aq. $NaHCO_3$, and diluted with CH_2Cl_2. The organic phase was separated and dried over $MgSO_4$, filtered, conc. *in vacuo*, and purified by flash column chromatography (hexane/EtOAc 6:1). Ethyl 2,3-di-*O*-benzyl-1-thio-β-D-glucopyranoside was obtained as a white solid (1.759 g, 92%). $\mathbf{R_f}$(hexane/EtOAc 1:1) 0.39. $[\boldsymbol{\alpha}]_{\boldsymbol{D}}^{\boldsymbol{298\,K}}$ = -24.3 (*c* 1, $CHCl_3$).[119] **^{1}H NMR** (400 MHz, $CDCl_3$) δ 7.41 – 7.29 (m, 10H, Ar*H*), 4.98 – 4.94 (m, 2H, OCH_2Ph), 4.77 – 4.73 (m, 2H, OCH_2Ph), 4.52 (d, $J_{1,2}$ = 9.7 Hz, 1H, H1), 3.87 (dd, $J_{6,6'}$ = 11.9 Hz, $J_{6,5}$ = 3.4 Hz, 1H, H6), 3.76 (dd, $J_{6',6}$ = 12.0 Hz, $J_{6',5}$ = 4.9 Hz, 1H, H6'), 3.60 (t, *J* = 9.2 Hz, 1H, H4), 3.51 (t, *J* = 8.8 Hz, 1H, H3), 3.42 – 3.38 (m, 1H, H2), 3.33 (ddd, $J_{5,4}$ = 9.3 Hz, $J_{5,6'}$ = 4.9 Hz, $J_{5,6}$ = 3.5 Hz, 1H, H5), 2.84 – 2.70 (m, 2H, SCH_2CH$_3$), 2.49 (s, 2H, OH), 1.33 (t, *J* = 7.4 Hz, 3H, SCH$_2$CH_3). **^{13}C NMR** (101 MHz, $CDCl_3$) δ 138.5 (1C, Ar*C*), 137.9 (1C, Ar*C*), 128.7 (2C, Ar*C*H), 128.5 (2C, Ar*C*H), 128.4 (2C, Ar*C*H), 128.1 (1C, Ar*C*H), 128.0 (1C, Ar*C*H), 128.0 (2C, Ar*C*H), 86.0 (C3), 85.4 (C1), 81.5 (C2), 79.2 (C5), 75.5 (O*C*H$_2$Ph), 75.5 (O*C*H$_2$Ph), 70.5 (C4), 62.7 (C6), 25.3 (S*C*H$_2$CH$_3$), 15.2 (SCH$_2$*C*H$_3$). **HRMS:** calcd. for $C_{22}H_{28}O_5SH^+$ *m/z* 405.1730; found *m/z* 405.1736. NMR data are in accordance with literature values.[119,178] To a solution of ethyl 2,3-di-*O*-benzyl-1-thio-β-D-glucopyranoside (1.717 g, 4.25 mmol, 1.0 equiv.) in dry pyridine (9.40 mL) was added TrtCl (1.302 g, 4.67 mmol, 1.1 equiv.). The mixture was stirred at 55 °C for 22 h at which point TLC (heptane/EtOAc 1:1) showed full conversion. The mixture was diluted with toluene and co-evaporated twice. Purified by flash column chromatography (heptane/EtOAc 7:1) to obtain the product **147** as a white foam (2.565 g, 93%). $\mathbf{R_f}$(hexane/EtOAc 2:1) 0.54. $[\boldsymbol{\alpha}]_{\boldsymbol{D}}^{\boldsymbol{298\,K}}$ = -27.4 (*c* 1, $CHCl_3$). **^{1}H NMR** (400 MHz, $CDCl_3$) δ 7.43 – 7.40 (m, 5H, Ar*H*), 7.38 – 7.17 (m, 20H, Ar*H*), 4.89 (d, *J* = 10.2 Hz, 1H, C2-OCH_2Ph), 4.86 (d, *J* = 11.3 Hz, 1H, C3-OCH_2Ph), 4.76 (d, *J* = 11.3 Hz, 1H, C3-OCH_2Ph), 4.71 (d, *J* = 10.3 Hz, 1H, C2-OCH_2Ph), 4.46 (d, *J* = 9.4 Hz, 1H, H1), 3.61 (td, *J* = 8.8, 2.2 Hz, 1H, H4), 3.46 (t, *J* = 8.7 Hz, 1H, H3), 3.39 (t, *J* = 9.1 Hz, 1H, H2), 3.37 – 3.31 (m, 3H, H5, H6, H6'), 2.85 – 2.66 (m, 2H, SCH_2CH$_3$), 2.43 (d, *J* = 2.4 Hz, 1H, O*H*), 1.32 (t, *J* = 7.4 Hz, 3H, SCH$_2$CH_3). **^{13}C NMR** (101 MHz, $CDCl_3$) δ 143.8 (3C, Ar*C*), 138.7 (1C, Ar*C*), 138.1 (1C, Ar*C*), 128.8 (6C, Ar*C*H), 128.7 (2C, Ar*C*H), 128.5 (2C, Ar*C*H), 128.5 (2C, Ar*C*H), 128.1 (2C, Ar*C*H), 128.1 (1C, Ar*C*H), 128.0 (6C, Ar*C*H), 128.0 (1C, Ar*C*H), 127.2 (3C, Ar*C*H), 87.1 (O*C*(Ph)$_3$), 86.2 (C3), 84.9 (C1), 81.4 (C2), 78.1 (C5), 75.7, 75.6 (O*C*H$_2$Ph), 72.4 (C4), 64.5 (C6), 24.8 (S*C*H$_2$CH$_3$), 15.4 (SCH$_2$*C*H$_3$). **HRMS:** calcd. for $C_{41}H_{42}O_5SNa^+$ *m/z* 669.2645; found *m/z* 669.2665. NMR data are in accordance with literature values.[76]

Ethyl 2,3-di-*O*-benzyl-4-*O*-methyl-1-thio-β-D-glucopyranoside (148)



To a solution of NaH (0.095 g, 2.36 mmol, 1.5 equiv.) in dry DMF (5.60 mL) at 0 °C was added a solution of **147** (0.996 g, 1.54 mmol, 1.0 equiv.) in dry DMF (5.60 mL). The mixture was stirred at 0 °C for 1 h, at which point MeI (0.14 mL, 2.25 mmol, 1.5 equiv.) dissolved in dry DMF (5.60 mL) was added and the reaction mixture was allowed to attain r.t. After 1.5 h TLC (hexane/EtOAc 3:1) showed full conversion. MeOH (6.0 mL) was slowly added to quench the reaction. Toluene was added, and the organic phase was washed with H_2O, dried over Na_2SO_4, filtered, and conc. *in vacuo*. The crude product was directly dissolved in $CHCl_3$/MeOH 2:1 (30 mL) and *p*TSA·H_2O was added until pH = 2. After 2 h TLC (hexane/EtOAc 3:1) showed full conversion. CH_2Cl_2 was added and the organic phase was washed with H_2O and $NaHCO_3$, dried over $MgSO_4$, filtered, and conc. *in vacuo*. The residue was purified by flash column chromatography (hexane/EtOAc 3:1), obtaining the product as a white solid (0.565 g, 88%). $\mathbf{R_f}$(hexane/EtOAc 3:1) 0.46. $[\boldsymbol{\alpha}]_{\boldsymbol{D}}^{\boldsymbol{298\,K}}$ = +22.9 (*c* 1, $CHCl_3$). **^{1}H NMR** (400 MHz, $CHCl_3$) δ 7.39 – 7.29 (m, 10H, Ar*H*), 4.91 – 4.83 (m, 3H, OCH_2Ph), 4.74 (d, J = 10.2 Hz, 1H, OCH_2Ph), 4.48 (d, $J_{1,2}$ = 9.8 Hz, 1H, H1), 3.90 (dd, $J_{6,6'}$ = 11.9 Hz, $J_{6,5}$ = 1.9 Hz, 1H, H6), 3.73 (dd, $J_{6',6}$ = 11.9 Hz, $J_{6',5}$ = 4.2 Hz, 1H, H6'), 3.62 – 3.58 (m, 1H, H3), 3.56 (s, 3H, OCH_3), 3.36 (dd, $J_{2,1}$ = 9.7 Hz, $J_{2,3}$ = 9.0 Hz, 1H, H2), 3.30 – 3.28 (m, 2H, H4, H5), 2.82 – 2.68 (m, 2H, SCH_2CH$_3$), 1.32 (t, J = 7.4 Hz, 3H, SCH$_2$CH_3). **^{13}C NMR** (101 MHz, $CHCl_3$) δ 138.6 (1C, Ar*C*), 138.1 (1C, Ar*C*), 128.6 (2C, Ar*C*H), 128.5 (2C, Ar*C*H), 128.4 (2C, Ar*C*H), 128.0 (3C, Ar*C*H), 127.9 (1C, Ar*C*H), 86.4 (C3), 85.3 (C1), 81.7 (C2), 80.0 (C4), 79.4 (C5), 75.8 (O*C*H$_2$Ph), 75.7 (O*C*H$_2$Ph), 62.3 (C6), 61.0 (O*C*H$_3$), 25.3 (SCH$_2$CH$_3$), 15.3 (SCH$_2$*C*H$_3$). **HRMS:** calcd. for $C_{23}H_{30}O_5SNa^+$ *m/z* 441.1706; found *m/z* 441.1704. NMR data are in accordance with literature values.[76]

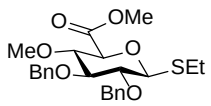
Ethyl 2,3-di-*O*-benzyl-4-*O*-methyl-1-thio-β-D-glucopyranosiduronic acid (149)



Primary alcohol **148** (0.627 g, 1.50 mmol, 1.0 equiv.) was dissolved in a solvent mixture of CH_2Cl_2 and H_2O 2:1 (6.6 mL) and stirred vigorously. TEMPO (0.045 g, 0.30 mmol, 0.2 equiv.) and $PhI(OAc)_2$ (1.208 g, 3.75 mmol, 2.5 equiv.) were added. Full conversion of the starting material was seen after 35 min with TLC (hexane/EtOAc 1:1, 1% AcOH). The remaining oxidant was quenched by addition of 10% aq. $Na_2S_2O_3$ (6.0 mL). The water phase was extracted with EtOAc, and the organic phase was dried over Na_2SO_4, filtered, and conc. *in vacuo*. The carboxylic acid **149** (0.527 g, 81 %) was obtained by

purification by flash column chromatography (heptane/EtOAc 2:1, 1% AcOH). $\mathbf{R_f}$(hexane/EtOAc 1:1, 1% AcOH) 0.44. $[\boldsymbol{\alpha}]_{\boldsymbol{D}}^{\boldsymbol{298\ K}}$ = -12.9 (*c* 1, $CHCl_3$). **^{1}H NMR** (400 MHz, $CHCl_3$) δ 7.37 – 7.29 (m, 10H, Ar*H*), 4.89 (d, *J* = 10.4 Hz, 1H, C2-OC*H*$_2$Ph), 4.85 (d, *J* = 11.0 Hz, 1H, C3-OC*H*$_2$Ph), 4.80 (d, *J* = 11.0 Hz, 1H, C3-OC*H*$_2$Ph), 4.72 (d, *J* = 10.4 Hz, 1H, C2-OC*H*$_2$Ph), 4.55 (d, $J_{1,2}$ = 9.7 Hz, 1H, H1), 3.88 (d, $J_{5,4}$ = 9.1 Hz, 1H, H5), 3.64 (t, *J* = 8.4 Hz, 1H, H3), 3.56 (d, *J* = 8.9 Hz, 1H, H4), 3.55 (s, 3H, OC*H*$_3$) 3.44 (dd, $J_{2,1}$ = 9.6 Hz, $J_{2,3}$ = 8.2 Hz, 1H, H2), 2.83 – 2.69 (m, 2H, SC*H*$_2$CH$_3$), 1.32 (t, *J* = 7.4 Hz, 3H, SCH$_2$C*H*$_3$). **^{13}C NMR** (101 MHz, $CHCl_3$) δ 172.3 (1C, *C*OOH), 138.2 (1C, Ar*C*), 137.8 (1C, Ar*C*), 128.6 (2C, Ar*C*H), 128.6 (2C, Ar*C*H), 128.4 (2C, Ar*C*H), 128.1 (1C, Ar*C*H), 128.0 (2C, Ar*C*H), 128.0 (1C, Ar*C*H), 85.7 (C1), 85.2 (C3), 80.9, 80.9 (C2, C4), 77.2 (C5), 75.7 (O*C*H$_2$Ph), 75.6 (O*C*H$_2$Ph), 60.9 (O*C*H$_3$), 25.5 (S*C*H$_2$CH$_3$), 15.1 (SCH$_2$*C*H$_3$). **HRMS:** calcd. for $C_{23}H_{28}O_6SNa^+$ *m/z* 455.1499; found *m/z* 455.1500.

Ethyl 2,3-di-*O*-benzyl-4-*O*-methyl-1-thio-β-D-glucopyranosiduronic acid (138)

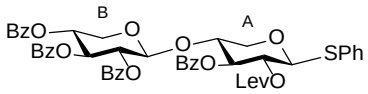


Carboxylic acid **149** (0.526 g, 1.22 mmol, 1.0 equiv.) was dissolved in a solvent mixture of dry toluene and MeOH 3.5:1 (5.4 mL) and $(CH_3)_3SiCHN_2$ 2 M in hexane (0.77 mL, 4.86 mmol, 4.0 equiv.) was slowly added. After 3 h TLC (hexane/EtOAc 2:1, 1% AcOH) showed full conversion of the starting material, and the excess reagent was quenched by addition of AcOH (2.0 mL). The mixture was conc. *in vacuo* and the crude material purified by flash column chromatography (pentane/EtOAc 10:1) affording the methyl ester **138** as a white amorphous solid (0.499 g, 92%). $\mathbf{R_f}$(pentane/EtOAc 10:1) 0.40. $[\boldsymbol{\alpha}]_{\boldsymbol{D}}^{\boldsymbol{298\ K}}$ = -2.9 (*c* 1, $CHCl_3$). **^{1}H NMR** (400 MHz, $CHCl_3$) δ 7.38 – 7.29 (m, 10H, Ar*H*), 4.91 – 4.86 (m, 2H, C2-OC*H*$_2$Ph, C3-OC*H*$_2$Ph), 4.83 (d, *J* = 11.0 Hz, 1H, C3-OC*H*$_2$Ph), 4.73 (d, *J* = 10.2 Hz, 1H, C2-OC*H*$_2$Ph), 4.48 (d, $J_{1,2}$ = 9.7 Hz, 1H, H1), 3.81 (s, 3H, COOC*H*$_3$), 3.79 (d, $J_{5,4}$ = 9.3 Hz, 1H, H5), 3.62 – 3.57 (m, 1H, H3), 3.55 (t, *J* = 8.9 Hz, 1H, H4), 3.51 (s, 3H, OC*H*$_3$), 3.43 (dd, $J_{2,1}$ = 9.6 Hz, $J_{2,3}$ = 8.6 Hz, 1H, H2), 2.82 – 2.67 (m, 2H, SC*H*$_2$CH$_3$), 1.31 (t, *J* = 7.4 Hz, 3H, SCH$_2$C*H*$_3$). **^{13}C NMR** (101 MHz, $CHCl_3$) δ 168.9 (1C, *C*OO), 138.4 (1C, Ar*C*), 138.0 (1C, Ar*C*), 128.5 (2C, Ar*C*H), 128.5 (2C, Ar*C*H), 128.4 (2C, Ar*C*H), 128.0 (1C, Ar*C*H), 128.0 (2C, Ar*C*H), 127.9 (1C, Ar*C*H), 85.9 (C1), 85.8 (C3), 81.3 (C4), 81.1 (C2), 78.0 (C5), 75.9 (C3-O*C*H$_2$Ph), 75.7 (C2-O*C*H$_2$Ph), 60.8 (O*C*H$_3$), 52.7 (COO*C*H$_3$), 25.2 (S*C*H$_2$CH$_3$), 15.1 (SCH$_2$*C*H$_3$). **HRMS:** calcd. for $C_{24}H_{30}O_6SNa^+$ *m/z* 469.1655; found *m/z* 469.1655. NMR data are in accordance with literature values.[76]

A: General procedure for glycosylation with a thiophenyl donor: Crushed molecular sieves (3 Å) is added a 2-necks Schleck flask, where the middle neck is fitted with a glass stopper, and the other with a septum. The flask is then placed under vacuum and heated. Afterwards an atm. of N_2 is applied. The donor (1.0 equiv.) dissolved in dry CH_2Cl_2 (c = 0.12 M) is added to the flask together with AgOTf (2.0 equiv.) dissolved in dry toluene (c = 0.36 M). Stirring of the reaction mixture is initiated and the glass stopper is exchanged for a thermometer under N_2 pressure, and the solution is cooled to -65 °C. 4-Nitrobenzene-sulfenyl chloride (1.0 equiv.) is dissolved in dry CH_2Cl_2 (c = 0.72 M) and slowly added as to not raise the temperature above -60 °C. The mixture is left to stir for approximately 10 min or until activation is complete. The acceptor (0.9 equiv.) is dissolved in dry CH_2Cl_2 (c = 0.65 M) and added quickly. The temperature is kept between -55 °C and -50 °C for an appropriate reaction time. Afterwards the mixture is allowed to heat to -15 °C over 10 min at which point Et_3N (3.0 equiv.) is added. The solution is filtered through a pad of Celite and conc. *in vacuo*. The residue is then purified by flash column chromatography to afford the product.

B: General procedure for glycosylation with a glucuronic acid donor: Crushed molecular sieves (3 Å) is added a 2-necks Schleck flask, where the middle neck is fitted with a glass stopper, and the other with a septum. The flask is placed under vacuum and heated. Afterwards an atm. of N_2 is applied. The donor (4.0 equiv.) dissolved in dry CH_2Cl_2 (c = 0.2 M) is added to the flask. Stirring of the reaction mixture is initiated and the glass stopper is exchanged for a thermometer under N_2 pressure, and the solution is cooled to -30 °C. Br_2 (1M in CH_2Cl_2) is slowly added until a yellow colour persists, then the mixture is left to stir for 10 min. Acceptor (1 equiv.) dissolved in dry CH_2Cl_2 (c = 0.2 M) is added alongside dry Et_2O (23 vol.%). $AgClO_4$ (4.8 equiv.) dissolved in dry toluene (c = 3.0 M) is added. The reaction mixture is stirred at -30 °C for 2 h, at which point Et_3N (5.0 equiv.) is added. The solution is allowed to attain r.t., filtered through a pad of Celite and conc. *in vacuo*. The residue is then purified by flash column chromatography to afford the product.

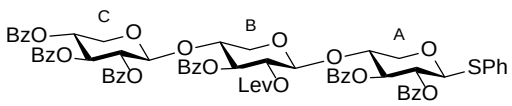
Phenyl 2,3,4-tri-*O*-benzoyl-β-D-xylopyranosyl-(1→4)-3-*O*-benzoyl-2-*O*-levulinoyl-1-thio-β-D-xylopyranoside (157)



General procedure A: Crushed molecular sieves (2.534 g). Donor **79** (1.019 g,1.84 mmol, 1.0 equiv.). AgOTf (0.944 g, 3.68 mmol, 2.0

equiv.). 4-Nitrobenzenesulfenyl chloride (0.350 g, 1.84 mmol, 1.0 equiv.). Acceptor **131** (0.721 g, 1.62 mmol, 0.9 equiv.). Reaction time 23 min. Et_3N (0.77 mL, 5.52 mmol, 3.0 equiv.). Eluent for flash column chromatography (heptane/toluene/EtOAc 6:2:2). The product was obtained as light yellow crystals (1.222 g, 85%). $\mathbf{R_f}$(hexane/toluene/EtOAc 3:1:2) 0.54. $[\boldsymbol{\alpha}]_{\boldsymbol{D}}^{\boldsymbol{298\,K}}$ = -43.6 (*c* 0.38, $CHCl_3$). **^{1}H NMR** (400 MHz, $CDCl_3$) δ 8.06 – 8.04 (m, 2H, Ar*H*), 7.97 – 7.91 (m, 6H, Ar*H*), 7.58 – 7.28 (m, 17H, Ar*H*), 5.64 (t, *J* = 6.6 Hz, 1H, $H3^B$), 5.50 (t, *J* = 8.2 Hz, 1H, $H3^A$), 5.22 (dd, $J_{2B,3B}$ = 6.6 Hz, $J_{2B,1B}$ = 4.9 Hz, 1H, $H2^B$), 5.11 – 5.04 (m, 2H, $H2^A$, $H4^B$), 4.91 (d, $J_{1B,2B}$ = 4.8 Hz, 1H, $H1^B$), 4.83 (d, $J_{1A,2A}$ = 8.6 Hz, 1H, $H1^A$), 4.19 (dd, $J_{5A,5A'}$ = 12.0 Hz, $J_{5A,4A}$ = 4.9 Hz, 1H, $H5^A$), 4.08 – 4.00 (m, 2H, $H4^A$, $H5^B$), 3.48 – 3.40 (m, 2H, H^5A', $H5^{B'}$), 2.64 – 2.46 (m, 4H, CH_2CH_2COO, $COCH_2CH_2$), 2.08 (s, 3H, CH_3CO). **^{13}C NMR** (101 MHz, $CDCl_3$) δ 205.9 (1C, C_q, CO, Lev), 171.4 (1C, C_q, COO, Lev), 165.6 (1C, C_q), 165.5 (1C, C_q), 165.3 (1C, C_q), 165.1 (1C, C_q) (CO, Bz), 133.5 (Ar*C*H), 133.5 (Ar*C*H), 133.5 (2C, Ar*C*H), 132.8 (2C, Ar*C*H), 132.5 (2C, Ar*C*H), 130.0 (1C, Ar*C*), 130.0 (2C, Ar*C*H), 129.9 (2C, Ar*C*H), 129.9 (2C, Ar*C*H), 129.5 (2C, Ar*C*H), 129.3 (1C, Ar*C*), 129.1 (1C, Ar*C*), 129.0 (2C, Ar*C*H), 128.6 (1C, Ar*C*), 128.6 (2C, Ar*C*H), 128.5 (2C, Ar*C*H), 128.5 (2C, Ar*C*H), 128.3 (Ar*C*H), 99.8 ($C1^B$), 86.5 ($C1^A$), 75.3 ($C4^A$), 73.5 ($C3^A$), 70.4 ($C2^A$), 70.2 ($C2^B$), 69.6 ($C3^B$), 68.6 ($C4^B$), 66.1 ($C5^A$), 60.9 ($C5^B$), 37.9 ($COCH_2CH_2$), 29.8 (CH_3CO), 28.1 (CH_2CH_2COO). **HRMS:** calcd. for $C_{49}H_{44}O_{14}SNa^+$ *m/z* 911.2343; found *m/z* 911.2362.

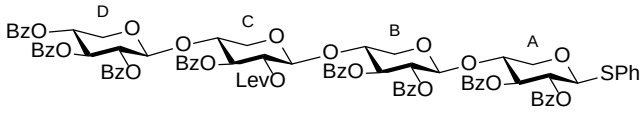
Phenyl 2,3,4-tri-*O*-benzoyl-β-D-xylopyranosyl-(1→4)-3-*O*-benzoyl-2-*O*-levulinoyl-β-D-xylopyranosyl-(1→4)-2,3-di-*O*-benzoyl-1-thio-β-D-xylopyranoside (158)



General procedure A: Crushed molecular sieves (0.3724 g). Donor **157** (1,131 g, 1.27 mmol, 1.0 equiv.). AgOTf (0.654 g, 2.55 mmol, 2.0 equiv.). 4-Nitrobenzenesulfenyl chloride (0.240 g, 1.27 mmol, 1.0 equiv.). Acceptor **80** (0.4158 g, 1.14 mmol, 0.9 equiv.). Reaction time 85 min. Et_3N (0.53 mL, 3.81 mmol, 3.0 equiv.). Eluent for flash column chromatography (hexane/toluene/EtOAc 2:1:1). The product was obtained as a white crystalline solid (0.768 g, 68%). $\mathbf{R_f}$(hexane/toluene/EtOAc 3:1:2) 0.46. $[\boldsymbol{\alpha}]_{\boldsymbol{D}}^{\boldsymbol{298\,K}}$ = -43.3 (*c* 0.24, $CHCl_3$). **^{1}H NMR** (400 MHz, $CDCl_3$) δ 8.00 – 7.88 (m, 12H, Ar*H*), 7.57 – 7.28 (m, 23H, Ar*H*), 5.62 – 5.56 (m, 2H, $H3^A$, $H3^C$), 5.35 (t, *J* = 8.4 Hz, 1H, $H3^B$), 5.32 (t, $J_{2A,1A}$ = 7.7 Hz, 1H, $H2^A$), 5.12 (dd, $J_{2C,3C}$ = 6.7 Hz, $J_{2C,1C}$ = 4.9 Hz, 1H, $H2^C$), 5.06 (d, $J_{1A,2A}$ = 7.8 Hz, 1H, $H1^A$), 5.06 – 5.02 (m, 1H, $H4^C$), 4.90 (dd, $J_{2B,3B}$ = 8.5 Hz, $J_{2B,1B}$ = 6.7 Hz, 1H, $H2^B$), 4.68 (d, $J_{1C,2C}$ = 4.8 Hz, 1H, $H1^C$), 4.59 (d, $J_{1B,2B}$ = 6.7 Hz, 1H, $H1^B$), 4.36 (dd, $J_{5Aeq,5Aax}$ = 12.2 Hz, $J_{5Aeq,4A}$ = 4.6 Hz, 1H, $H5^A{}_{eq}$), 4.01 – 3.95 (m, 2H,

H4^{A}, H5^{C}), 3.76 (td, J = 8.5 Hz, $J_{4B,5Beq}$ = 5.0 Hz, 1H, H4^{B}), 3.63 (dd, $J_{5Aax,5Aeq}$ = 12.1 Hz, $J_{5Aax,4A}$ = 8.3 Hz, 1H, H5$^{A}_{ax}$), 3.52 (dd, $J_{5Beq,5Bax}$ = 12.2 Hz, $J_{5Beq,4B}$ = 4.9 Hz, 1H, H5$^{B}_{eq}$), 3.35 (dd, J = 12.3, 6.2 Hz, 1H, H5$^{C'}$), 3.09 (dd, $J_{5Bax,5Beq}$ = 12.2 Hz, $J_{5Bax,5B4A}$ = 9.0 Hz, 1H, H5$^{B}_{ax}$), 2.65 – 2.36 (m, 4H, CH_2CH_2COO, $COCH_2CH_2$), 2.07 (s, 3H, CH_3CO). **^{13}C NMR** (101 MHz, $CDCl_3$) δ 206.0 (1C, C_q, CO, Lev), 171.3 (1C, C_q, COO, Lev), 165.6 (1C, C_q), 165.5 (2C, C_q), 165.4 (1C, C_q), 165.4 (1C, C_q), 165.3 (1C, C_q), 165.0 (1C, C_q) (CO, Bz), 133.6 (1C, ArCH), 133.6 (1C, ArCH), 133.5 (1C, ArCH), 133.5 (1C, ArCH), 133.3 (1C, ArCH), 133.2 (1C, ArCH), 133.0, 132.7 (2C, ArCH), 130.1 (2C, ArCH), 130.0 (2C, ArCH), 130.0 (2C, ArCH), 129.9 (2C, ArCH), 129.9 (2C, ArCH), 129.9 (2C, ArCH), 129.7 (1C, ArC), 129.5 (1C, ArC), 129.3 (1C, ArC), 129.2 (1C, ArC), 129.1 (2C, ArCH), 129.0 (1C, ArC), 128.6 (2C, ArCH), 128.6 (2C, ArCH), 128.6 (2C, ArCH), 128.5 (2C, ArCH), 128.5 (2C, ArCH), 128.4 (2C, ArCH), 128.1 (1C, ArCH), 100.8 (C1^{B}), 99.5 (C1^{C}), 86.7 (C1^{A}), 75.5 (C4^{A}), 75.0 (C4^{B}), 72.8 (C3^{A}), 72.5 (C3^{B}), 71.6 (C2^{B}), 70.5 (C2^{A}), 70.1 (C2^{C}), 69.6 (C3^{C}), 68.6 (C4^{C}), 65.6 (C5^{A}), 62.4 (C5^{B}), 60.9 (C5^{C}), 37.8 ($COCH_2CH_2$), 29.8 (CH_3CO), 27.9 (CH_2CH_2COO). **HRMS:** calcd. for $C_{68}H_{60}O_{20}SNa^+$ *m/z* 1251.3290; found *m/z* 1251.3311.

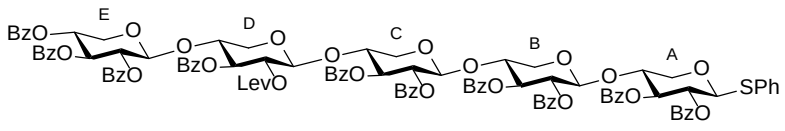
Phenyl 2,3,4-tri-*O*-benzoyl-β-D-xylopyranosyl-(1→4)-3-*O*-benzoyl-2-*O*-levulinoyl-β-D-xylopyranosyl-(1→4)-2,3-di-*O*-benzoyl-β-D-xylopyranosyl-(1→4)-2,3-di-*O*-benzoyl-1-thio-β-D-xylopyranoside (159)



General procedure A: Crushed molecular sieves (0.490 g). Donor **158** (0.201 g, 0.16 mmol, 1.0 equiv.). AgOTf (0.085 g, 0.33 mmol, 2.0 equiv.). 4-Nitrobenzenesulfenyl chloride (0.030 g, 0.16 mmol, 1.0 equiv.). Acceptor **80** (0.065 g, 0.14 mmol, 0.9 equiv.). Reaction time 90 min. Et_3N (0.07 mL, 0.48 mmol, 3.0 equiv.). Eluent for flash column chromatography (hexane/toluene/EtOAc 2:1:1). The product was obtained as a white crystalline solid (0.174 g, 77%). **R$_f$**(hexane/toluene/EtOAc 3:1:2) 0.38. $[\alpha]_D^{298\,K}$ = -42.1 (*c* 0.38, $CHCl_3$). **^{1}H NMR** (400 MHz, $CDCl_3$) δ 8.01 – 7.86 (m, 16H, ArH), 7.57 – 7.50 (m, 6H), 7.45 – 7.25 (m, 22H), 7.18 – 7.13 (m, 1H), 5.63 (t, J = 7.9 Hz, 1H, H3^{A}), 5.57 (t, J = 6.7 Hz, 1H, H3^{D}), 5.46 (t, J = 7.7 Hz, 1H, H3^{B}), 5.33 – 5.27 (m, 2H, H2^{A}, H3^{C}), 5.15 (dd, $J_{2B,3B}$ = 7.9 Hz, $J_{2B,1B}$ = 6.0 Hz, 1H, H2^{B}), 5.11 (dd, $J_{2D,3D}$ = 6.4 Hz, $J_{2D,1D}$ = 4.7 Hz, 1H, H2^{D}) 5.03 (td, J = 6.3, 4.1 Hz, 1H, H4^{D}), 4.98 (d, $J_{1A,2A}$ = 8.0 Hz, 1H, H1^{A}), 4.83 (dd, J = 8.6 Hz, $J_{2C,1C}$ = 6.8 Hz, 1H, H2^{C}), 4.78 (d, $J_{1B,2B}$ = 6.1 Hz, 1H, H1^{B}), 4.66 (d, $J_{1D,2D}$ = 4.9 Hz, 1H, H1^{D}), 4.35 (d, $J_{1C,2C}$ = 6.8 Hz, 1H, H1^{C}), 4.18 (dd, $J_{5A,5'A}$ = 12.1 Hz, $J_{5A,4A}$ = 4.7 Hz, 1H, H5$^{A}_{eq}$), 4.03 (td, J = 8.2 Hz, $J_{4A,5A}$ = 4.8 Hz, 1H, H4^{A}), 3.97

(dd, $J_{5Deq,5Dax}$ = 12.4 Hz, $J_{5Deq,4D}$ = 3.9 Hz, 1H, $H5^{D}_{eq}$), 3.78 – 3.65 (m, 2H, $H4^{B}$, $H5^{B}$, $H4^{C}$), 3.51 – 3.44 (m, 2H, $H5'^{A}$, $H5^{C}$), 3.33 (dd, $J_{5Dax,5Deq}$ = 12.4 Hz, $J_{5Dax,4D}$ = 6.3 Hz, 1H, $H5^{D}_{ax}$), 3.28 (dd, $J_{5Bax,5Beq}$ = 11.7 Hz, $J_{5Bax,4B}$ = 7.9 Hz, 1H, $H5^{B}_{ax}$), 3.01 (dd, $J_{5Cax,5Ceq}$ = 12.2 Hz, $J_{5Cax,4C}$ = 9.1 Hz, 1H, $H5^{C}_{ax}$), 2.61 – 2.46 (m, 2H, $COCH_2CH_2$), 2.44 – 2.30 (m, 2H, CH_2CH_2COO), 2.06 (s, 3H, CH_3CO). **^{13}C NMR** (101 MHz, $CDCl_3$) δ 205.8 (1C, C_q, CO, Lev), 171.2 (1C, C_q, COO, Lev), 165.6 (1C, C_q), 165.5 (2C, C_q), 165.4 (1C, C_q), 165.3 (1C, C_q), 165.3 (1C, C_q), 165.1 (1C, C_q), 165.0 (1C, C_q) (CO, Bz), 133.5 (2C, Ar*C*H), 133.5 (1C, Ar*C*H), 133.4 (1C, Ar*C*H), 133.4 (2C, Ar*C*H), 133.3 (1C, Ar*C*H), 133.1 (1C, Ar*C*H), 132.9 (1C, Ar*C*), 132.6 (2C, Ar*C*H), 130.1 (2C, Ar*C*H), 130.0 (2C, Ar*C*H), 130.0 (2C, Ar*C*H), 129.9 (2C, Ar*C*H), 129.9 (2C, Ar*C*H), 129.9 (2C, Ar*C*H), 129.8 (2C, Ar*C*H), 129.8 (2C, Ar*C*H), 129.7 (1C, Ar*C*), 129.6 (1C, Ar*C*), 129.5 (1C, Ar*C*), 129.5 (1C, Ar*C*), 129.4 (1C, Ar*C*), 129.3 (1C, Ar*C*), 129.1 (1C, Ar*C*), 129.1 (2C, Ar*C*H), 129.0 (1C, Ar*C*), 128.6 (4C, Ar*C*H), 128.5 (4C, Ar*C*H), 128.5 (2C, Ar*C*H), 128.5 (4C, Ar*C*H), 128.4 (2C, Ar*C*H), 128.1 (1C, Ar*C*H), 100.8 ($C1^{B}$), 100.6 ($C1^{C}$), 99.5 ($C1^{D}$), 86.7 ($C1^{A}$), 75.4 ($C4^{A}$), 75.3, 75.1 ($C4^{B}$, $C4^{C}$), 73.0 ($C3^{A}$), 72.5 ($C3^{C}$), 71.9 ($C3^{B}$), 71.4, 71.3 ($C2^{B}$, $C2^{C}$), 70.5 ($C2^{A}$), 70.1 ($C2^{D}$), 69.6 ($C3^{D}$), 68.6 ($C4^{D}$), 65.7 ($C5^{A}$), 62.4, 62.3 ($C5^{B}$, $C5^{C}$), 60.9 ($C5^{D}$), 37.8 ($COCH_2CH_2$), 29.8 (CH_3CO), 27.9 (CH_2CH_2COO). **HRMS:** calcd. for $C_{87}H_{76}O_{26}SNa^+$ *m/z* 1591.4237; found *m/z* 1591.4254.

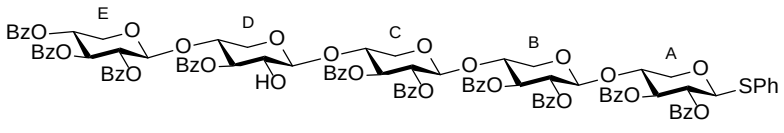
Benzyl 2,3,4-tri-*O*-benzoyl-β-D-xylopyranosyl-(1→4)-3-*O*-benzoyl-2-*O*-levulinoyl-β-D-xylopyranosyl-(1→4)-2,3-di-*O*-benzoyl-β-D-xylopyranosyl-(1→4)-2,3-di-*O*-benzoyl-β-D-xylopyranoside-(1→4)-2,3-di-*O*-benzoyl-β-D-xylopyranoside (150)



General procedure A: Crushed molecular sieves (0.754 g). Donor **159** (0.400 g, 0.26 mmol, 1.0 equiv.). AgOTf (0.1317 g, 0.51 mmol, 2.0 equiv.). 4-Nitrobenzenesulfenyl chloride (0.048 g, 0.26 mmol, 1.0 equiv.). Acceptor **130** (0.103 g, 0.23 mmol, 0.9 equiv.). Reaction time 90 min. Et_3N (0.11 mL, 0.77 mmol, 3.0 equiv.). Eluent for flash column chromatography (hexane/toluene/EtOAc 4:3:2). The product was obtained as a white crystalline solid (0.364 g, 83%). $\mathbf{R_f}$(hexane/toluene/EtOAc 3:3:2) 0.33. $[\boldsymbol{\alpha}]_{\boldsymbol{D}}^{\boldsymbol{298\ K}}$ = -43.0 (*c* 0.10, $CHCl_3$). **^{1}H NMR** (400 MHz, $CDCl_3$) δ 7.98 – 7.85 (m, 20H, Ar*H*), 7.57 – 7.49 (m, 6H, Ar*H*), 7.45 – 7.16 (m, 29H, Ar*H*), 5.57 (t, *J* = 6.7 Hz, 1H, $H3^{E}$), 5.55 – 5.51 (m, 1H, $H3^{A}$), 5.49 (t, *J* = 8.0 Hz, 1H, $H3^{B}$), 5.40 (t, *J* = 8.0 Hz, 1H, $H3^{C}$), 5.31 – 5.27 (m, 2H, $H2^{A}$, $H3^{D}$), 5.15 (dd, *J* = 8.4, 6.6 Hz, 1H, $H2^{B}$), 5.12 (dd, *J* = 6.9, 5.1 Hz, 1H, $H2^{E}$), 5.08 – 5.01 (m, 2H, $H2^{C}$, $H4^{E}$), 4.84 – 4.80 (m, 1H, $H2^{D}$), 4.80 (d, *J* = 12.5 Hz, 1H,

OC*H*$_2$Ph), 4.70 (d, *J* = 6.3 Hz, 1H, H1^B), 4.66 (d, *J* = 4.9 Hz, 1H, H1^E), 4.64 (d, *J* = 6.6 Hz, 1H, H1^A), 4.55 (d, *J* = 12.4 Hz, 1H, OC*H*$_2$Ph), 4.54 (d, *J* = 6.0 Hz, 1H, H1^C) 4.33 (d, *J* = 6.7 Hz, 1H, H1^D), 3.99 – 3.94 (m, 3H, H4^A, H5^A, H5^E), 3.80 – 3.73 (m, 2H, H4^B, H4^D), 3.69 (td, *J* = 8.0, 7.5, 4.4 Hz, 1H, H4^C), 3.60 (dd, *J* = 12.1, 4.8 Hz, 1H, H5^C), 3.48 – 3.43 (m, 2H, H5^B, H5^D), 3.40 – 3.31 (m, 2H, H5$^{A'}$, H5$^{E'}$), 3.20 (dd, *J* = 12.2, 8.4 Hz, 1H, H5$^{C'}$), 3.09 (dd, *J* = 12.2, 8.6 Hz, 1H, H5$^{B'}$), 3.00 (dd, *J* = 12.2, 9.0 Hz, 1H, H5$^{D'}$), 2.53 (q, *J* = 6.7 Hz, 2H, COC*H*$_2$CH$_2$), 2.38 – 2.33 (m, 2H, CH$_2$C*H*$_2$COO), 2.05 (s, 3H, C*H*$_3$CO). **^{13}C NMR** (101 MHz, $CDCl_3$) δ 205.7 (1C, C_q, CO, Lev), 171.2 (1C, C_q, COO, Lev), 165.6 (1C, C_q), 165.5 (1C, C_q), 165.4 (1C, C_q), 165.4 (1C, C_q), 165.4 (1C, C_q), 165.3 (2C, C_q), 165.1 (1C, C_q), 165.0 (1C, C_q), 165.0 (1C, C_q) (CO, Bz), 136.9 (1C, Ar*C*), 133.5 (2C, Ar*C*H), 133.5 (1C, Ar*C*H), 133.4 (1C, Ar*C*H), 133.4 (1C, Ar*C*H), 133.3 (1C, Ar*C*H), 133.2 (2C, Ar*C*H), 133.1 (1C, Ar*C*H), 133.1 (1C, Ar*C*H), 130.0 (2C, Ar*C*H), 130.0 (2C, Ar*C*H), 130.0 (2C, Ar*C*H), 129.9 (3C, Ar*C*H), 129.8 (4C, Ar*C*H), 129.8 (2C, Ar*C*H), 129.7 (3C, Ar*C*H), 129.7 (1C, Ar*C*), 129.7 (1C, Ar*C*), 129.6 (1C, Ar*C*), 129.6 (1C, Ar*C*), 129.5 (1C, Ar*C*), 129.4 (1C, Ar*C*), 129.4 (1C, Ar*C*), 129.3 (1C, Ar*C*), 129.1 (1C, Ar*C*), 129.0 (1C, Ar*C*), 128.6 (4C, Ar*C*H), 128.6 (2C, Ar*C*H), 128.5 (3C, Ar*C*H), 128.5 (3C, Ar*C*H), 128.5 (2C, Ar*C*H), 128.4 (3C, Ar*C*H), 128.4 (2C, Ar*C*H), 128.4 (2C, Ar*C*H), 128.3 (2C, Ar*C*H), 127.9 (1C, Ar*C*H), 127.9 (3C, Ar*C*H), 101.1 (C1^B), 100.6, 100.5 (C1^C, C1^D), 99.5, 99.4 (C1^A, C1^E), 76.1 (C4^A), 75.3, 75.2, 75.0 (C4^B, C4^C, C4^D), 72.5 (C3^D), 72.2, 72.1, 72.0 (C3^A, C3^B,C3^E), 71.5, 71.4 (C2^B, C2^D), 71.2 (C2^C), 71.1 (C2^A), 70.4 (O*C*H$_2$Ph), 70.1 (C2^E), 69.6 (C3^E), 68.6 (C4^E), 62.7 (C5^A), 62.4, 62.3 (2C) (C5^B, C5^C, C5^D), 60.9 (C5^E), 37.8 (CO*C*H$_2$CH$_2$), 29.7 (*C*H$_3$CO), 27.9 (CH$_2$*C*H$_2$COO). **HRMS:** calcd. for $C_{107}H_{94}O_{33}Na^+$ *m/z* 1930.5603; found *m/z* 1930.5619.

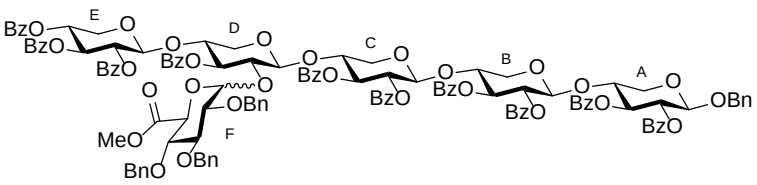
Benzyl 2,3,4-tri-*O*-benzoyl-β-D-xylopyranosyl-(1→4)-3-*O*-benzoyl-β-D-xylopyranosyl-(1→4)-2,3-di-*O*-benzoyl-β-D-xylopyranosyl-(1→4)-2,3-di-*O*-benzoyl-β-D-xylopyranoside-(1→4)-2,3-di-*O*-benzoyl-β-D-xylopyranoside (151)



Pentasaccharide **150** (0.324 g, 0.17 mmol, 1.0 equiv.) was dissolved in pyridine/AcOH 2:1 (0.68 mL). A mixture of 50% solution of hydrazine hydrate (20 µL, 0.34 mmol, 2.0 equiv.) dissolved in pyridine (c = 20 M) was added. The reaction mixture was left to stir at r.t. overnight, at which point TLC showed consumption of the starting material and formation of a product. The reaction was stopped by addition of acetone (3.7 mL, 51.0 mmol, 300 equiv.) and was left to stir at r.t. for 3 h. EtOAc was added, and the mixture washed with 1 M HCl and brine, filtered, and

conc. *in vacuo*. The residue was purified by flash column chromatography (hexane/toluene/EtOAc 3:3:2) affording the product as a white powder (0.249 g, 81%). $\mathbf{R_f}$(hexane/toluene/EtOAc 3:3:2) 0.34. $[\boldsymbol{\alpha}]_{\boldsymbol{D}}^{\boldsymbol{298\ K}}$ = -38.3 (*c* 1, $CHCl_3$). **^{1}H NMR** (400 MHz, $CDCl_3$) δ 8.02 – 7.84 (m, 20H, Ar*H*), 7.59 – 7.30 (m, 30H, Ar*H*), 7.22 – 7.16 (m, 5H, Ar*H*), 5.64 (t, *J* = 6.7 Hz, 1H, $H3^E$), 5.52 (t, *J* = 8.1 Hz, 1H, $H3^A$), 5.49 (t, *J* = 8.1 Hz, 1H, $H3^B$), 5.40 (t, *J* = 7.9 Hz, 1H, $H3^C$), 5.29 (dd, *J* = 8.5, 6.7 Hz, 1H, $H2^A$), 5.21 (t, *J* = 7.4 Hz, 1H, $H3^D$), 5.18 – 5.09 (m, 3H, $H2^B$, $H2^E$, $H4^E$), 5.06 (dd, *J* = 8.1, 6.1 Hz, 1H, $H2^C$), 4.80 (d, *J* = 12.3 Hz, 1H, OC*H*$_2$PH), 4.72 (d, *J* = 4.9 Hz, 1H, $H1^E$), 4.70 (d, *J* = 6.4 Hz, 1H, $H1^B$), 4.63 (d, *J* = 6.6 Hz, 1H, $H1^A$), 4.54 (d, *J* = 5.3 Hz, 1H, $H1^C$), 4.55 (d, *J* = 12.5 Hz, 1H, OC*H*$_2$PH), 4.27 (d, *J* = 5.6 Hz, 1H, $H1^D$), 4.15 (dd, *J* = 12.4, 3.9 Hz, 1H, $H5^E$), 4.00 – 3.94 (m, 2H, $H4^A$, $H5^A$), 3.79 – 3.70 (m, 3H, $H4^B$, $H4^C$, $H4^D$), 3.63 – 3.57 (m, 2H, $H5^C$, $H5^D$), 3.47 – 3.43 (m, 3H, $H5^B$, $H2^D$, $H5'^E$), 3.37 (td, *J* = 10.1, 3.5 Hz, 1H, $H5'^A$), 3.14 (dd, *J* = 12.0, 7.7 Hz, 1H, $H5'^C$), 3.12 – 3.03 (m, 2H, $H5'^B$, $H5'^D$). **^{13}C NMR** (101 MHz, $CDCl_3$) δ 166.1 (1C, C_q), 165.6 (1C, C_q), 165.5 (2C, C_q), 165.5 (1C, C_q), 165.4 (1C, C_q), 165.3 (1C, C_q), 165.3 (1C, C_q), 165.1 (1C, C_q), 165.0 (1C, C_q) (CO, Bz), 136.9 (1C, Ar*C*), 133.6 (Ar*C*H), 133.6 (1C, Ar*C*H), 133.5 (2C, Ar*C*H), 133.4 (1C, Ar*C*H), 133.4 (1C, Ar*C*H), 133.3 (1C, Ar*C*H), 133.3 (1C, Ar*C*H), 133.2 (1C, Ar*C*H), 133.1 (1C, Ar*C*H), 130.1 (2C, Ar*C*H), 130.0 (3C, Ar*C*H), 130.0 (3C, Ar*C*H), 129.9 (2C, Ar*C*H), 129.9 (4C, Ar*C*H), 129.8 (3C, Ar*C*H), 129.8 (4C, Ar*C*H), 129.8 (2C, Ar*C*H), 129.7 (1C, Ar*C*H), 129.6 (1C, Ar*C*), 129.6 (1C, Ar*C*), 129.6 (1C, Ar*C*), 129.5 (1C, Ar*C*), 129.4 (3C, Ar*C*), 129.3 (1C, Ar*C*), 129.2 (1C, Ar*C*), 129.0 (1C, Ar*C*), 128.6 (3C, Ar*C*H), 128.6 (2C, Ar*C*H), 128.5 (5C, Ar*C*H), 128.4 (3C, Ar*C*H), 128.4 (2C, Ar*C*H), 128.3 (2C, Ar*C*H), 127.9 (1C, Ar*C*H), 127.9 (3C, Ar*C*H), 101.7 ($C1^D$, $J_{C\text{-}H}$ = 164 Hz), 101.2 ($C1^B$, $J_{C\text{-}H}$ = 162 Hz), 100.6 ($C1^C$, $J_{C\text{-}H}$ = 162 Hz), 99.4 ($C1^A$, $J_{C\text{-}H}$ = 163 Hz), 99.0 ($C1^E$, $J_{C\text{-}H}$ = 165 Hz), 76.2 ($C4^A$), 75.4 ($C4^B$), 74.0 ($C4^C$, $C4^D$), 73.2 ($C3^D$), 72.3, 72.2 ($C3^A$, $C3^B$), 71.8 ($C3^C$), 71.6 ($C2^B$), 71.2, 71.1 (2C) ($C2^A$, $C2^C$, $C2^D$), 70.4 (OCH_2Ph), 70.2 ($C2^E$), 69.5 ($C3^E$), 68.6 ($C4^E$), 62.7 ($C5^A$), 62.4 ($C5^B$), 62.0 ($C5^C$), 61.3($C5^D$), 61.0 ($C5^E$). **HRMS:** calcd. for $C_{102}H_{88}O_{31}Na^+$ *m/z* 1831.5202; found *m/z* 1831.5112.

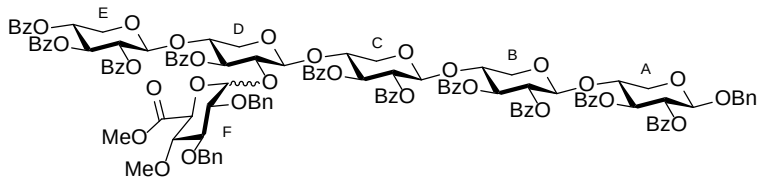
Benzyl 2,3,4-tri-*O*-benzoyl-β-D-xylopyranosyl-(1→4)-[methyl (1→2)-2,3,4-tri-*O*-benzyl-α/β-D-glucuronatyl]-3-*O*-benzoyl-β-D-xylopyranosyl-(1→4)-2,3-di-*O*-benzoyl-β-D-xylopyranosyl-(1→4)-2,3-di-*O*-benzoyl-β-D-xylopyranoside-(1→4)-2,3-di-*O*-benzoyl-β-D-xylopyranoside (152)



General procedure B: Crushed molecular sieves (0.395 g). Donor **137** (0.281 g, 0.54 mmol, 4.0 equiv.). Acceptor **151** (0.243 g, 0.13 mmol, 1.0 equiv.). $AgClO_4$ (0.133 g, 0.64 mmol, 4.8 equiv.). Et_3N (0.09 mL, 0.67 mmol, 5.0 equiv.). Eluent for flash column chromatography (elution of product: heptane/toluene/acetone 4:1:1, elution of acceptor: heptane/toluene/acetone 2:1:1). The product was obtained as a white crystalline anomeric mixture (α/β 4:1) (0.138 g, 45%), and acceptor **151** was furthermore reisolated (0.057 g, 23%). **R_f**(heptane/toluene/acetone 2:1:2) 0.50. **^{1}H NMR** (400 MHz, $CDCl_3$) δ 8.03 – 7.82 (m, 28H, Ar*H*), 7.59 – 7.15 (m, 75H, Ar*H*), 7.01 – 6.98 (m, 2H, Ar*H*), 5.57 – 5.24 (m, 9H, H2^A, H3^A, H2^B, H3^C, H3^D, H3^E, H1^F), 5.12 (dd, *J* = 8.0, 6.4 Hz, 1H, H3^B), 5.08 (dd, *J* = 6.8, 5.0 Hz, 1H, H2^E), 4.97 (td, *J* = 6.4, 4.2 Hz, 1H, H4^E), 4.92 (dd, *J* = 7.3, 5.5 Hz, 1H, H2^C), 4.85 – 4.52 (m, 14H, H1^A, H1^B, H1^E, C1^{A}OC*H*$_2$Ph), 4.44 (d, *J* = 3.4 Hz, 0.25H), 4.39 (d, *J* = 5.4 Hz, 0.25H), 4.34 (d, *J* = 11.2 Hz, 1H), 4.22 (d, *J* = 7.2 Hz, 1H, H1^D), 4.17 (d, *J* = 5.4 Hz, 1H, H1^C), 4.08 (d, *J* = 10.0 Hz, 1H), 4.01 – 3.93 (m, 2.5H, H4^A, H5^A, C1^{A}OC*H*$_2$Ph), 3.89 – 3.79 (m, 3H, H5^E), 3.72 – 3.58 (m, 5.5H, H4^B, H2^D, H4^D), 3.56 – 3.44 (m, 7H, H5^B, H4^C, H5^C, H5^D), 3.40 (s, 3H, OC*H*$_3$), 3.37 – 3.33 (m, 1H, H5$^{A'}$), 3.26 – 3.20 (m, 2H, H5$^{E'}$), 3.11 – 2.97 (m, 2H, H5^B', H5$^{C'}$), 2.96 – 2.91 (m, 1H, H5$^{D'}$). **^{13}C NMR** (101 MHz, $CDCl_{3f}$) δ 170.2 (1C, C_q, *C*OOCH$_3$), 169.0 ($C_β$), 165.7 (1C, C_q), 165.5 (1C, C_q), 165.4 (2C, C_q), 165.4 (1C, C_q), 165.4 (1C, C_q), 165.3 (1C, C_q), 165.1 (1C, C_q), 165.0 (1C, C_q), 165.0 (1C, C_q) (CO, Bz), 138.6, 138.4, 138.3, 136.9, 133.5, 133.5, 133.4, 133.3, 133.2, 133.2, 133.1, 133.1, 130.2, 130.1, 130.1, 130.0, 129.9, 129.9, 129.8, 129.8, 129.6, 129.4, 129.3, 129.2, 129.0, 128.6, 128.6, 128.5, 128.4, 128.4, 128.4, 128.3, 128.2, 128.1, 128.1, 128.0, 127.9, 127.9, 127.8, 127.5, 127.4 (Ar*C*, Ar*C*H), 102.2 (C1^D), 101.0 (C1^B), 99.8, 99.5, 99.3 (C1^A, C1^C, C1^E), 97.3 (C1^F, $J_{C\text{-}H}$ = 173 Hz), 80.7 (C3^F), 80.0 (C4^F), 79.3 (C2^F), 76.6 (C2^D), 76.2, 76.1 (C4^A, C4^D), 75.8 (C3^{F}O*C*H$_2$Ph), 74.5 (C4^{F}O*C*H$_2$Ph), 74.1 (C4^B), 73.7 (C3^D), 73.5 (C2^{F}O*C*H$_2$Ph), 73.4 (C4^C), 72.4 (C3^A), 71.8 (C3^B), 71.4 (C2^B), 71.2, 71.1, 70.8 (C2^A, C2^C, C3^C), 70.5 (C5^F), 70.4 (C1^{A}O*C*H$_2$Ph), 70.3 (C2^E), 69.7 (C3^E), 68.7 (C4^E), 63.1, 62.8 (C5^A, C5^D), 61.9 (C5^B), 61.4 (C5^C), 60.9 (C5^E), 52.6 ($C_β$), 52.1 (COO*C*H$_3$).

Only the α-product has been fully characterised. However for the β-anomer some signals can be detected in the ^{1}H and ^{13}C NMR spectra which have been reported above.

Benzyl 2,3,4-tri-*O*-benzoyl-β-D-xylopyranosyl-(1→4)-[methyl (1→2)-2,3-di-*O*-benzyl-4-*O*-methyl-α/β-D-glucuronatyl]-3-*O*-benzoyl-β-D-xylopyranosyl-(1→4)-2,3-di-*O*-benzoyl-β-D-xylopyranosyl-(1→4)-2,3-di-*O*-benzoyl-β-D-xylopyranoside-(1→4)-2,3-di-*O*-benzoyl-β-D-xylopyranoside (153)



General procedure B: Crushed molecular sieves (0.372 g). Donor **138** (0.197 g, 0.44 mmol, 4.0 equiv.). Acceptor **151** (0.200 g, 0.11 mmol, 1.0 equiv.). $AgClO_4$ (0.111 g, 0.53 mmol, 4.8 equiv.). Et_3N (0.08 mL, 0.55 mmol, 5.0 equiv.). Eluent for flash column chromatography (heptane/toluene/acetone 6:3:2). The product was obtained as a white crystalline anomeric mixture (α/β 2.5:1) (0.106 g, 44%), and acceptor **151** was furthermore reisolated (0.034 g, 17%). $\mathbf{R_f}$(heptane/toluene/acetone 2:1:2) 0.45. **^{1}H NMR** (400 MHz, $CDCl_3$) δ 8.04 – 7.83 (m, 33H, Ar*H*), 7.59 – 7.11 (m, 79H, Ar*H*), 5.58 – 5.24 (m, 10.7H, H2^{A}, H3^{A}, H3^{B}, H3^{C}, H3^{D}, H3^{E}, H1^{F}), 5.17 – 5.06 (m, 3H, H2^{B}, H2^{E}), 5.04 (dd, *J* = 7.0, 5.1 Hz, 0.4H), 5.00 – 4.93 (m, 1.6H, H4^{E}), 4.90 (dd, *J* = 7.5, 5.6 Hz, 1H, H2^{C}), 4.83 – 4.82 (m, 2.8H, C1^{A}OC*H*$_2$Ph, C3^{F}OC*H*$_2$Ph), 4.79 – 4.73 (m, 2H), 4.71 – 4.68 (m, 2.4H, H1^{B}, C2^{F}OC*H*$_2$Ph), 4.66 – 4.60 (m, 3.7H, H1A, H1^{E}, C3^{F}OC*H*$_2$Ph), 4.57 – 4.52 (m, 3.3H, C1^{A}OC*H*$_2$Ph, C2^{F}OC*H*$_2$Ph), 4.50 (d, *J* = 5.1 Hz, 0.4H), 4.45 (d, *J* = 11.5 Hz, 0.4H), 4.39 (d, *J* = 5.4 Hz, 0.4H), 4.21 (d, *J* = 7.1 Hz, 1H, H1^{D}), 4.13 (d, *J* = 5.6 Hz, 1H, H1^{C}), 4.02 – 3.95 (m, 4H, H4^{A}, H5^{A}, H5^{F}), 3.89 – 3.56 (m, 11.3H, H4^{B}, H2^{D}, H4^{D}, H5^{E}, H3^{F}), 3.53 – 3.34 (m, 14H, H5$^{A'}$, H5^{B}, H4^{C}, H5^{C}, H5^{D}, H2^{F}, COOC*H*$_3$), 3.29 – 3.18 (m, 7.4H, H5$^{E'}$, H4^{F}, OC*H*$_3$), 3.11 – 3.02 (m, 2H, H5$^{B'}$), 2.99 (dd, *J* = 11.4, 6.8 Hz, 1H, H5$^{C'}$), 2.95 – 2.89 (m, 1H, H5$^{D'}$). **^{13}C NMR** (101 MHz, $CDCl_3$) δ 170.3 (1C, C_q, *C*OOCH$_3$), 169.1 (C_q), 165.7 (1C, C_q), 165.5 (1C, C_q), 165.4 (1C, C_q), 165.4 (1C, C_q), 165.4 (1C, C_q), 165.3 (1C, C_q), 165.3 (1C, C_q), 165.3 ($C_β$), 165.2 ($C_β$), 165.1 (1C, C_q), 165.0 ($C_β$), 164.9 (1C, C_q), 164.9 (1C, C_q) (*C*O, Bz) 138.7, 138.6 ($C_β$), 138.5, 136.9, 133.5, 133.5, 133.4, 133.3, 133.2, 133.2, 133.1, 133.0, 130.2, 130.1, 130.1, 130.1, 123.0, 129.9, 129.9, 129.8, 129.8, 129.7 ($C_β$), 129.6 ($C_β$), 129.6, 129.6, 129.5 ($C_β$), 129.4 ($C_β$), 129.4, 129.3, 129.3 ($C_β$), 129.2, 129.0, 128.6, 128.5, 128.4, 128.4, 128.4, 128.3, 128.2, 128.2, 128.1, 128.0, 127.9, 127.9, 127.8, 127.7 ($C_β$), 127.6 ($C_β$), 127.4 (Ar*C*, Ar*C*H), 103.1 ($C_β$), 102.0 (C1^{D}), 101.1

(C1^{B}), 100.9 (C_{β}), 99.9 (C1^{E}), 99.6 (C_{β}), 99.5 (2C, C1^{A}, C1^{C}), 99.4 (C_{β}), 97.6 (C1^{F}, $J_{C\text{-}H}$ = 173 Hz), 83.6 (C_{β}), 81.7 (C4^{F}), 80.6 (C3^{F}), 79.1 (C2^{F}), 76.8 (C2^{D}), 76.6 (C4^{D}), 76.1 (C4^{A}), 76.1 (C_{β}), 75.7 (C3^{F}O*C*H$_2$Ph), 75.5 (C_{β}), 75.4 (C_{β}), 75.1 (C_{β}), 75.1 (C_{β}), 74.8 (C_{β}), 74.5 (C_{β}), 74.2 (C4^{B}), 74.1 (C_{β}), 73.7 (C3^{D}), 73.5 (C2^{F}O*C*H$_2$Ph), 73.5 (C4^{C}), 72.6 (C_{β}), 72.4 (C3^{A}), 72.3 (C_{β}), 71.8 (C3^{B}), 71.5 (C_{β}), 71.4, 71.2, 71.2 (C2^{A}, C2^{B}, C3^{C}), 71.1 (C_{β}), 70.6, 70.5 (C2^{C}, C2^{E}), 70.4 (C1^{A}O*C*H$_2$Ph), 70.3 (C5^{F}), 70.2 (C_{β}), 69.8 (C3^{E}), 68.7 (C4^{E}), 62.9, 62.8 (C5^{A}, C5^{D}), 62.7 (C_{β}), 62.0 (C5^{B}), 61.5 (C5^{C}), 60.9 (C5^{E}), 60.7 (C_{β}), 60.1 (O*C*H$_3$), 52.7 (C_{β}), 52.2 (COO*C*H$_3$).

Only the α-product has been fully characterised. However for the β-anomer some signals can be detected in the ^{1}H and ^{13}C NMR spectra which have been reported above.

4. Evaluation of ferulic acid esterase activity

The content of this chapter covers the research project on ferulic acid esterase (FAE) activity. The results were obtained during a three months external stay at the Food Chemistry Department of Wageningen University and Research, The Netherlands, under the supervision of Associate Professor Mirjam Kabel and with the guidance of Dr. Matthias Frommhagen.

The main focus of this study was to expand the knowledge of FAEs as a tool in biodegradation through investigation of the reactivity and specificity of several FAEs towards both synthetic model and natural substrates. The tested FAEs were provided by Westerdijk Fungal Biodiversity Institute, Utrecht University, Utrecht, The Netherlands, and the Food Chemistry Department of Wageningen University and Research, Wageningen, The Netherlands.

The first part of this chapter will introduce the FAEs investigated in this study as well as the substrates used. The obtained results will then be presented and finally these results will be discussed and concluding remarks provided.

4.1 Aim of the project

The aim of this product was to investigate the activity and reactivity of 14 fungal FAEs covering six different subfamilies (SF) (as according to the 2016 classification[27]) towards both model and natural substrates. Hence, in this project the aim was to expand the knowledge of the specificity of these FAEs in releasing hydroxycinnamic acids from natural plant substrates such

as xylan (-like), pectin and lignin. Thereby gaining insight into how the reactivity correlates with the recent phylogenetic analysis dividing the FAEs into 13 SFs[27] compared to the ABCD classification.[27] Furthermore, the investigation can possible identify an enzyme or enzymes which could have potential commercial value for the degradation of biomass materials e.g. by addition to enzyme mixtures.

4.2 Ferulic acid esterases investigated in this project

This project was made possible by the donation of 11 fungal FAEs by Westerdijk Fungal Biodiversity Institute and three FAEs from the Food Chemistry Department. Together these 14 enzymes cover six SFs (SF1, 5, 6, 7, 9, 13) (Table 4-1). The 11 enzymes received from the Westerdijk Fungal Biodiversity Institute were produced as active enzymes by heterologous expression using *P. pastoris* as the host organism,[66] and the reactivity of these 11 enzymes towards the four common model substrates have previously been investigated.[66] The enzymes were provided as the culture filtrates from the growth. The three enzymes provided by the Food Chemistry Department at Wageningen University and Research were purified enzymes, which have previously been characterised on both synthetic model substrates and some natural substrates.[179] Generally, the enzymes will be referred to by their name hereafter. Furthermore, a sample of the *P. pastoris* without insertion of FAE genes (named “broth”) were provided for negative control.

Table 4-1: List over the enzymes investigated. The fungal origin, accession numbers, names used for this research as well as SF classification and former ABCD classification is provided.[66,179,180]

Fungi	Accession number	Name	SF classification[a]	ABCD classification[b]
Aspergillus niger	CAC83933.1	AnFaeB	1	B
Aspergillus sydowii	Aspsy1_293049	AsFaeF	1	B
Aspergillus niger	XP_001395336.1	AnFaeC	5	C
Aspergillus nidulans	AN5267	AnidFAEC	5	C or D
Aspergillus sydowii	Aspsy1_154482	AsFaeC	5	C or D
Chrysosporium lucknowense C1[c]	JF826027	C1FaeA1	5	A
Chrysosporium lucknowense C1[c]	JF826028	C1FaeA2	5	A
Fusarium oxysporum	Fusox1_8990	FoFae2	6	n/a
Aspergillus sydowii	Aspsy1_1158585	AsFaeE	6	C or D
Chrysosporium lucknowense C1[c]	JF826029	C1FaeB2	6	B
Aspergillus niger	CAA70510.1	AnFaeA	7	A
Aspergillus niger	An15g05280	AnFaeJ	9	n/a
Aspergillus sydowii	Aspsy1_160668	AsFaeI	13	B
Stereum hirsutum	Stehi1_73641	ShFae1	13	n/a

[a]SF classification as according to the phylogenetic tree made in 2016.[27] [b]Classification assigned according to previous summation by Dilokpimol *et al.* in 2016[27] and Antonopoulou *et al.* in 2018.[180] [c]Purified enzymes provided by the Food Chemistry Department. n/a: not applicable.

The activity of the three enzymes provided by the Food Chemistry Department have previously been tested towards several insoluble cell wall polysaccharides (wheat bran, Brewer's spent grain, corn cobs, corn fibre arabinoxylan, and sugar beet pulp) but found inactive. However, after pre-treatment of the polysaccharides to form soluble oligomers with different xylanases or commercial enzyme preparations, the FAEs proved active. C1FaeA1 was active in different degree towards all substrates, but especially active against wheat bran, and able to release both di- and triFAs. C1FaeA2 was generally less active, the main activity was towards corn fibre arabinoxylan oligomers, but it was able to release triFA to a greater extent than diFA. C1FaeB2 was most active on Brewer's spent grain and SBP oligomers but releases only small amounts of di- and triFAs.[179]

4.3 Natural substrates tested

Six different natural compounds were tested as substrates for the FAEs. These were chosen to represent different classes of plant material: xylan (-like), pectin, and lignin. Hence, the ability of the FAEs to release monomeric FA and pCA as well as polymeric hydroxycinnaminic acids

linked to different sugar units could be evaluated. The non-commercial natural substrates were provided by the Food Chemistry Department at Wageningen University and Research.

4.3.1 Sugar beet pectin

The plant cell wall in roots of various dicots are rich in pectin, therefore commercial sugar beet pectin (SBP) was utilised as a representative for pectin. Hence, possible FAE activity would indicate that hydroxycinnamic acids linked to the neutral sidechains arabinosyl (*O*-2 and *O*-5 linked) and galactosyl (*O*-6 linked) could be cleaved.

4.3.2 Corn fiber oligomers

In the corn-processing industry an important and abundant by-product is corn fiber, and this by-product comprises approximately 10% of the processed dry corn, of which 40% is hemicellulose.[181] The hemicellulose in corn fiber mainly consists of GAX. The use of corn fiber as a source for bioethanol production would be of great value, but the enzymatic hydrolysis has to be improved. The structural features of GAX from corn fiber have previously been studied at the Food Chemistry Department by M. Appeldoorn *et al.*[182,183] Corn fiber was subjected to pre-treatment with mild diluted acid, enzymatic hydrolysis by hemicellulases and fermentation, which removed fermentable monosaccharides of glucose, xylose, and arabinose. Then an apolar fraction was separated by solid phase column, which represented 26% (w/w) of the dry matter and 39% (w/w) of the total soluble oligosaccharides in the starting material. This fraction called corn fiber oligomers (CFoligo) was highly ferulated, and subsequent analysis concluded that almost all oligomers in the material contained FA or diFA substituents,[183] which due to the GAX nature of the substrate must be *O*-5 linked to Ara*f*.

The CFoligo fraction resisted hydrolysis of commercial enzymes available in 2013 and as described above had a high content of FA and diFAs,[183] and it therefore represented a very interesting material for the investigation of specificity of the FAEs in this project.

4.3.3 Insoluble wheat arabinoxylans

Commercial insoluble wheat arabinoxylans (WAX-i) as a substrate was used as a general representative for insoluble xylan, such as found in plant cell wall materials rich in xylan (e.g. corn fiber, and various grasses).[25,184] Due to the xylan nature the hydroxycinnamic acids would be ester-linked to *O*-5 at Ara*f* units.

4.3.4 Corn stover

Corn stover is defined as the above-ground, non-grain portion of the crop including stem, cob and leaves. This is a large and available source of biomass for biofuel production, which would not compete with food production. The United States alone have an estimated corn stover production of 200 to 250 million dry tons per year.[185,186] This substrate was included as it represents both a ready resource, but also a more natural lignocellulosic mixture of lignin, cellulose, and hemicellulose.[187]

The corn stover material investigated as a substrate for the FAEs in this project came from corn stover which had been freeze-dried and extracted with acetone, after which insoluble material was filtered off, dried, and ball milled to a fine powder.[188]

4.3.5 Wheat straw

The definition of wheat straw are all the harvestable wheat residues after the grains have been collected. This biomass resource like the others also have potential for biofuel production. Like corn stover wheat straw also represent a more natural lignocellulosic mixture of compounds, i.e. cellulose, hemicellulose, and lignin (37.8%, 26.5%, and 17.5%, respectively).[189]

Like for corn stover the fine powder of wheat straw used as a substrate for this project was obtained by freeze-drying the wheat straw, the material was then extracted with acetone, the insoluble material isolated, and ball milled.[188]

4.3.6 Corn stover lignin

Pure lignin isolates from corn stover (CSlignin) have previously been isolated in the laboratory at the Food Chemistry Department.[188] Lignin as a FAE substrate is interesting as this is an aromatic polymer with incorporation of FA, diFAs and more so of pCA, which is mostly reported to be γ-esterified to β-*O*-4-linkages,[32] but is often assumed inaccesable to FAEs.

4.4 Investigation of FAE specificity and reactivity

4.4.1 Determination of protein content in enzyme filtrates

The protein content of the enzyme filtrates from Westerdijk Fungal Biodiversity Institute had previously been approximated by a densitometry method, but a more accurate measurement of the protein content was required. The received filtrates were slightly yellow in colour, and the enzymes were therefore first repeatedly filtered and the medium exchanged for a sodium acetate buffer, after which the protein content was measured using bovine serum albumin

(BSA) as a standard in a spectrophotometric assay.[190] The standard covered a range of 2 – 2000 µg/mL, which was chosen based on the approximated protein content received with the enzymes. However, when the absorbance of the enzymes were measured it was found to be outside the range of the standard curve. Hence, all the samples were diluted twice with MilliQ to be within the range of the standards. A linear relationship with an uncertainty of maximum 10% between the diluted and undiluted samples was assumed as dilution of the standards showed linearity with an uncertainty of maximum 9%. The calculated protein contents can be seen in Table 4-2.

Table 4-2: Calculated protein content in the filtered and concentrated enzyme samples from Westerdijk Fungal Biodiversity Institute.

Enzyme	Protein content (mg/mL)	Enzyme	Protein content (mg/mL)
AnFaeB	33.3	AsFaeE	28.4
AsFaeF	27.0	AnFaeA	34.0
AnFaeC	27.4	AnFaeJ	26.5
AnidFAEC	32.5	AsFaeI	26.1
AsFaeC	27.7	ShFae1	26.3
FoFae2	28.2	Broth	3.4

The calculated protein content in the samples were as mentioned higher than expected. To exclude the possibility of interference from reducing agents, which could be present in the broth from the growth medium, a spectrophotometric assay which is compatible with reducing agents[191] was performed on the broth from *P. pastoris* without insertion. The calculated protein content of the broth from *P. pastoris* without insertion from both spectrophotometric assays were comparable (3.8 and 3.4 mg/mL), and therefore interference from reducing agents could be excluded. The broth without insertion has not been used as a negative blank for this experiment as the exact treatment of the enzymes prior to arrival to the Food Chemistry Department is not known, and consequently there could be a difference in the degree of concentration or other unknown factors. As a conclusion, the protein contents measured in Table 4-2 was assumed accurate. Based on the results all the enzymes were diluted to 1 mg/mL by addition of sodium acetate buffer. Hence, a more precise dosing with a 10% insecurity of the enzymes can be performed for further experiments.

4.4.2 FAE activity towards synthetic substrates

The two synthetic model substrates methyl ferulate (MF) and methyl *p*-coumarate (MpC) were chosen as substrates among the four most common synthetic esters (MF, MpC, methyl caffeate,

and methyl sinapate) used when testing for FAE activity, as these represent the major hydroxycinnamic acids present in the plant cell walls.

The most commonly used screening method for FAE activity is a spectrophotometric method.[27,63,66,179,192] This type of assay is based on the de-methylation of an aromatic methyl ester to the corresponding aromatic acid, which is fast and easy to perform. Even though the spectrophotometric assay is widely used and accepted there is concern about the spectral overlapping of the substrates and the products as their absorbance maxima are quite close (Table 4-3).[27]

Table 4-3: Absorbance maxima of the synthetic substrates used in this assay and their corresponding acids.[193]

Phenolic compound	λ_{max} (pH 6.0)	Phenolic compound	λ_{max} (pH 6.0)
Methyl ferulate	321.5	Methyl *p*-coumarate	310
Ferulic acid	310	*p*-Coumaric acid	285

Consequently, all quantifications in this study were performed using reversed phase ultra-high performance liquid chromatography (RP-UHPLC) coupled to a ultraviolet (UV) detector and electron spray ionization (ESI) mass spectroscometry (MS) for more precise measurements. The use of UHPLC compared to HPLC allowed for improved resolution or maintaining the same resolution while decreasing the run time due to the utilization of a smaller particle size of the column. The small particles give a better resolution as they provide a more uniform flow through the column, and furthermore the distance the solute must diffuse in the mobile and stationary phase is in order of the particle size. A small particle size together with a narrow column and a higher flow means that the analyte does not get very diluted which lowers the detection limit. However, this leads to higher resistance to solvent flow meaning that a higher pressure is required which is possible using UHPLC.[194]

The activity of the 14 FAEs towards the two model substrates MF and MpC were determined after 2 h and 19 h of incubation, and all reactions were done in duplicates. The results were quantified by UHPLC-UV, providing the results as absolute values, where a reduction in substrate concentration was considered as FAE activity (starting concentration was 2 mM) (Figure 4-1, for numerical values and standard deviations see Table A1 in Appendix).

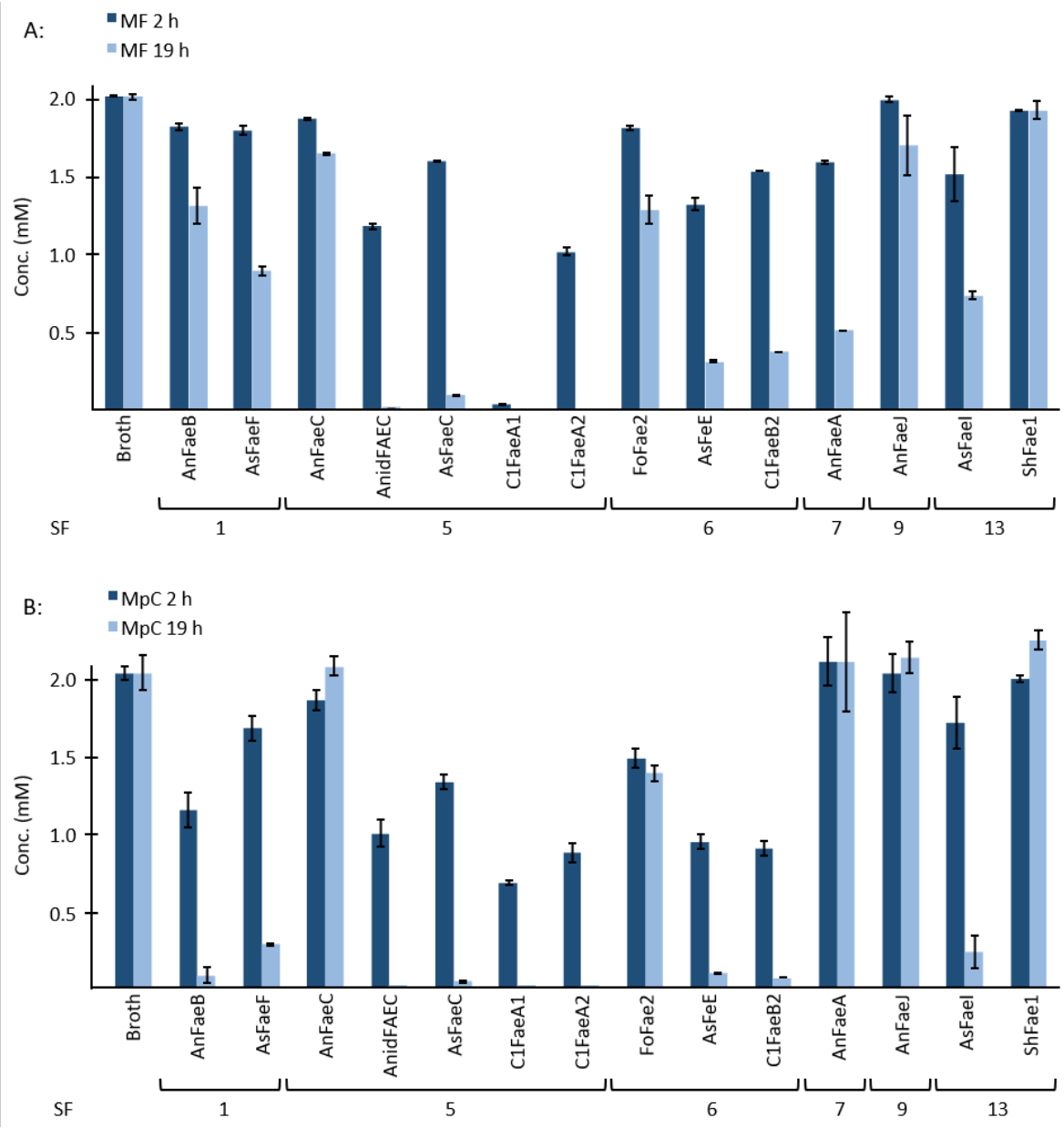


Figure 4-1: FAE reactivity towards the model substrates MF (A) and MpC (B) measured as reduction in concentration from a starting concentration of 2 mM.

Interference from the broth could be excluded as no decrease in the MF and MpC concentration was observed, this moreover confirmed, that no interfering enzymes were introduced by the production organism. Furthermore, it is seen that the standard deviation between the two duplicates for all entries was quite low indicating good reproducibility.

From Figure 4-1 it is seen that incubation of MF with all the FAEs except AnFaeJ (SF9) and ShFae1 (SF13) showed a decrease in MF concentration after 2 h, which was even more pronounced after 19 h. Generally, the enzymes belonging to SF5 were able to de-methylate MF to a high degree at the enzyme dose used, but AnFaeC (SF5) did not follow the same trend.

Furthermore, the enzymes AnFaeC (SF5), AnFaeA (SF7), AnFaeJ (SF9), and ShFae1 (SF13) did not prove able to hydrolyse MpC after 2 h or 19 h, when these FAEs were incubated with MpC. For the other 9 FAEs incubation with MpC led to a decrease in MpC concentration after 2 h to different extends, and after 19 h incubation nearly all MpC had been hydrolysed.

4.4.3 Type and quantification of hydroxycinnamic acids from the natural substrates

The free and ester-bound amounts of FA, pCA, five diFAs (8-8'-(furan)-diFA, 8-8'-aryl-diFA, 8-5'-diFA, 5-5'-diFA, 8-O-4'-diFA (Figure 1-7)), two structurally unknown triFAs and two compounds of m/z = 401 from the six different natural plant cell wall substrates were analysed.

The free and ester-bound content of the quantified compounds were determined by saponification, which were done in triplicates and analysed after 19 h incubation. The free content was determined from the substrate samples without addition of enzyme or broth. The ester-bound content was determined as the difference between the total and free content (Table 4-4).

For the obtained FA and pCA standard curves, both the UV and MS signals from the RP-UHPLC-UV-ESI-MS analysis showed linear dependency on the FA and pCA concentration, the MS curves resulted in R^2 between 0.9723 and 0.9997 (no further data shown). Consequently, for CFoligo, corn stover, wheat straw, and CSlignin the quantification of di- and triFAs were done by MS. The quantification for SBP and WAX-i were done using UV due to low amounts of released compounds. Generally was the contents of di- and triFAs calculated relative to FA and mass corrected. Furthermore, due to the lower substrate availability of CSlignin, all results for this substrate generally display higher standard deviations (e.g. see Table 4-4).

Generally, the distribution of FA, pCA, diFAs, and triFAs differed greatly depending on the substrate. As an example, CFoligo was highly ferulated with monomeric compounds, diFAs, and triFAs, but triFAs were not found in corn stover, wheat straw nor CSlignin. However, CSlignin contained a high amount of ester-bound pCA.

Table 4-4: Free and ester-bound content in μg/mg sample of the quantified mono, di- and triFAs in the six natural substrates tested.

Compound	Type	SBP[a]	CFoligo	WAX-i[a]	Corn stover	Wheat straw	CSlignin
FA[b]	Free	0.05±0.01	74.54±1.55	0.02.00	0.13±0.00	0.06±0.00	0.01±0.00
	Bound	15.04±0.31	57.32±1.15	15.04±0.15	12.71±0.63	5.51±0.09	8.82±0.16
pCA[b]	Free	0.04±0.00	5.58±0.20	0.04±0.00	1.13±0.01	0.39±0.00	0.78±0.01
	Bound	0.21±0.00	6.54±0.12	0.98±0.04	28.39±1.37	6.60±0.16	98.25±2.79
8-8'-(furan)-diFA	Free	0.02±0.00[c]	1.08±0.12	0.014±0.00[c]	0.46±0.05[c]	1.01±0.08[c]	0.026±0.00[c]
	Bound	0.60±0.12[c]	19.96±0.41	0.51±0.03[c]	27.86±1.04[c]	30.29±0.63[c]	45.93±2.20[c]
8-8'-aryl-diFA	Free	0	0	0	0	0	0
	Bound	0	45.12±2.32	1.13±0.04	3.39±0.23	0.99±0.05	2.37±0.17
8-5'-diFA	Free	0	4.13±0.15	0	0	0	0
	Bound	0.23±0.01	56.63±1.20	1.13±0.04	13.83±0.66	2.45±0.09	5.45±0.23
m/z = 401 $_{9.74\ min}$	Free	0	7.10±0.03	0	0.20±0.00	0	0
	Bound	0	19.90±0.43	0	5.54±0.08	0.97±0.02	9.31±0.04
5-5'-diFA	Free	0	15.59±1.02	0	0	0	0
	Bound	0.21±0.00	153.83±4.16	2.58±0.07	8.35±0.27	2.66±0.09	1.29±0.13
m/z = 401 $_{10.66\ min}$[d]	Free	0	5.83±0.25	0	0	0.17±0.00	0
	Bound	0.04±0.00	n/a	0.18±0.01	0.32±0.01	0.16±0.02	0
8-O-4'-diFA	Free	0	3.55±0.10	0	0	0	0
	Bound	0.54±0.02	105.51±2.61	2.97±0.00	6.00±0.16	2.54±0.19	0
TriFA 1 $_{11.35\ min}$	Free	0	0.64±0.10	0	0	0	0
	Bound	0.06±0.00	26.76±0.93	2.16±0.06	0	0	0
TriFA 2 $_{12.38\ min}$	Free	0	0	0	0	0	0
	Bound	0.05±0.00	13.86±0.71	2.22±0.061	0	0	0

All numbers represent μg/mg sample. [a]Quantification based on UV. [b]Quantified by UV which show linear correlation to MS quantification. [c]Major mass peak: m/z = 389. Former mass peak: m/z = 403. Could correspond to a reduced form (*vide infra*). [d]Measured bound amount less than what has been observed released in some instances, hence the compound is rendered unstable under the basic conditions used for saponification. n/a: not applicable. "Compound$_{min}$" indicate the compound and retention time in subscript.

As an example for the elution profiles obtained by the UHPLC analysis the profile of CFoligo prior to and after incubation with the two FAEs C1FaeA1 and AnFaeA can be seen in Figure 4-2.

For CFoligo, FA was both the most abundant compound of those quantified as well as it was highly released upon incubation with the two shown enzymes (C1FaeA1 and AnFaeA) (Figure 4-2A). From Figure 4-2 it can furthermore be seen that several diFAs were released, and of these the 5-5'diFA was the most pronounced.

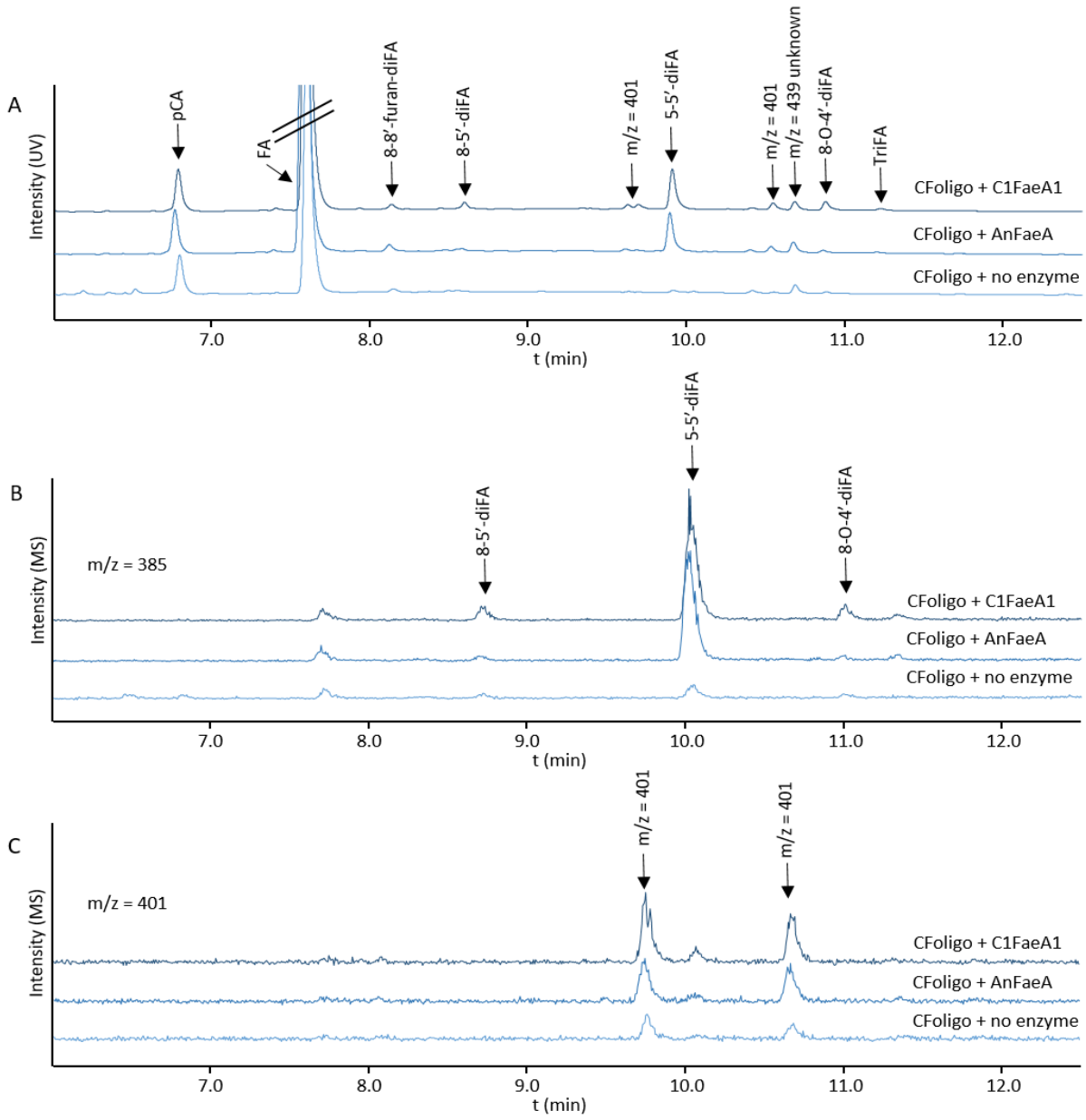


Figure 4-2: UHPLC elution profiles for CFoligo without added enzyme (corresponding to free content), and CFoligo incubated with C1FaeA1 or AnFaeA. A) UV chromatograms, B) Elution profiles for m/z = 385 (diFAs), C) Elution profiles for m/z = 401. B and C were subtracted from the MS data. Masses represent negative ions; loss of 1 Da.

The diFAs and one of triFAs (R_t 11.35 min) quantified in this study have been characterised by retention time, MS, and MS^2, where MS^2 indicate tandem mass spectrometry (Table 4-5, for the MS^2 spectra see Figure A1-A9 in Appendix). For the triFA eluting at 12.28 min it was not possible to obtain a MS^2 spectrum, probably due to limited released amounts.

Table 4-5: Retention time, calculated mass, observed mass, and MS^2 fragmentation of the quantified di- and triFA compounds.

Compound	Retention time$_{UV}$ (min)	Exact calculated mass	Observed mass $[M-H]^{-a}$	MS^2 data (intensity)[b]
8-8'-(furan)-diFA	8.23	404.1107	403	134 (7.8), 151 (7.2), 165 (7.1), 178 (7.5), 193 (100), 209 (5.0), 215 (12.9), 341 (59.9)
8-8'-(furan)-diFA reduced form	8.22	390.1315	389	151 (5.8), 165 (7.0), 178 (6.4), 181 (7.8), 193 (100), 195 (16.0), 321 (30), 341 96.9), 343 (13.0)
8-8'-(aryl)-diFA	8.36	386.1002	385	341 (100)
8-5'-diFA	8.72	386.1002	385	297 (29.6), 341 (100)
$m/z = 401_{9.74\ min}$	9.74	402.0951	401	325 (8.0), 357 (100.0)
5-5'-diFA	10.03	386.1002	385	282 (6.9), 326 (11.6), 341 (100), 342 (9.7), 370 (13.5)
$m/z = 401_{10.66\ min}$	10.66	402.0951	401	191 (25.3), 235 (9.1), 357 (30.9), 371 (34.5), 383 (100)
8-O-4'-diFA	11.00	386.1002	385	193 (100), 249 (8.1), 313 (70.4), 317 (7.4), 341 (79.1)
TriFA 1	11.34	578.1424	577	311 (6.5), 355 (34.2), 489 (5.7), 533 (100.0), 534 (6.2)

[a]$[M-H]^-$ refers to the observed mass peak representing the negative ions; loss of 1 Da. [b]Intensity: the highest peak correspond to 100%, the other peaks are given as percentages hereof. Peaks of more than 5% have been reported.

The peak which eluted at 8.3 min (Figure 4-2A) for CFoligo had a mass of m/z = 403 (404 Da) corresponding to the 8-8'-(furan)-diFA, for which the identity and presence in the plant cell wall have previously been confirmed both by synthesis[195,196] and observation.[182] Interestingly, during the study it was observed, that for the other substrates the retention time of the peak

did not change, but in MS the predominant peak had a mass of m/z = 389 (390 Da). This mass could correspond to a reduced form of the diFA (Figure 4-3). The theory can possibly further be corroborated by the MS^2 spectrum from the fragmentation pattern (Table 4-5 and Figure A2 in Appendix).

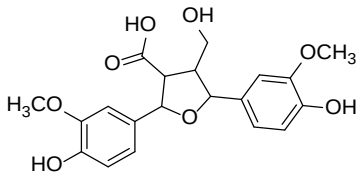


Figure 4-3: Proposed structure for m/z 389 (390 Da) as a reduced form of 8-8'-(furan)-diFA.

The identity of the other diFAs 8-8'-(aryl), 8-5', 5-5', and 8-O-4' have been confirmed by previous reported data[31] as well as former results obtained at the Food Chemistry Department (unpublished results).

As mentioned above many of the FAEs were found to be able to release two unknown compounds of m/z = 401 (402 Da) (R_t 9.74 and 10.66 min) (Figure 4-2 C). The release of a compound with this mass have also previously been observed.[179,182] Based on the mass it is speculated that these compounds can be dehydrodimers composed of FA and 5-hydroxyferulic acid, where the second is known to be involved in the hydroxycinnamic biosynthesis.[197,198] However it was observed, that $m/z = 401_{10.66\ min}$ seemed to be unstable under the saponification conditions used, which was concluded by the observation that in several instances the released amounts were greater than the measured ester-bound (Table 4-4 and Figure 4-6 *vide infra*). Based on differences in elution time, stability under saponification, and MS^2 fragmentation the two compounds with m/z = 401 must be regioisomers (Figure 4-2, Table 4-5, Figure A5 and A7).

4.4.4 FAE activity towards natural substrates

The additional release of the aforementioned hydroxycinnamic compounds from the six natural substrates due to FAE activity were then determined, which resulted in an increase of the free compounds. All enzyme reactions were incubated for 19 h and were done in duplicates. The determined free content of the individual hydroxycinnamic compound of the specific substrate was used as a negative blank, and afterwards the released amount was related to the maximum ester-bound content as percentages hereof.

The release of FA and pCA by the 14 FAEs and the broth were quantified by UHPLC-UV for all incubations with the natural substrates (Figure 4-4, the numerical values and standard deviations can be seen in Table A2 in Appendix).

The results showed that even though pCA was released to a lesser extent than FA, both compounds were released from CFoligo, corn stover, wheat straw, and CSlignin. However, only very low amounts of FA and pCA were released from SBP and WAX-i, while release for the broth was absent. Furthermore, it can be observed, that the enzymes categorised as SF5, 6, and 7 were the most active.

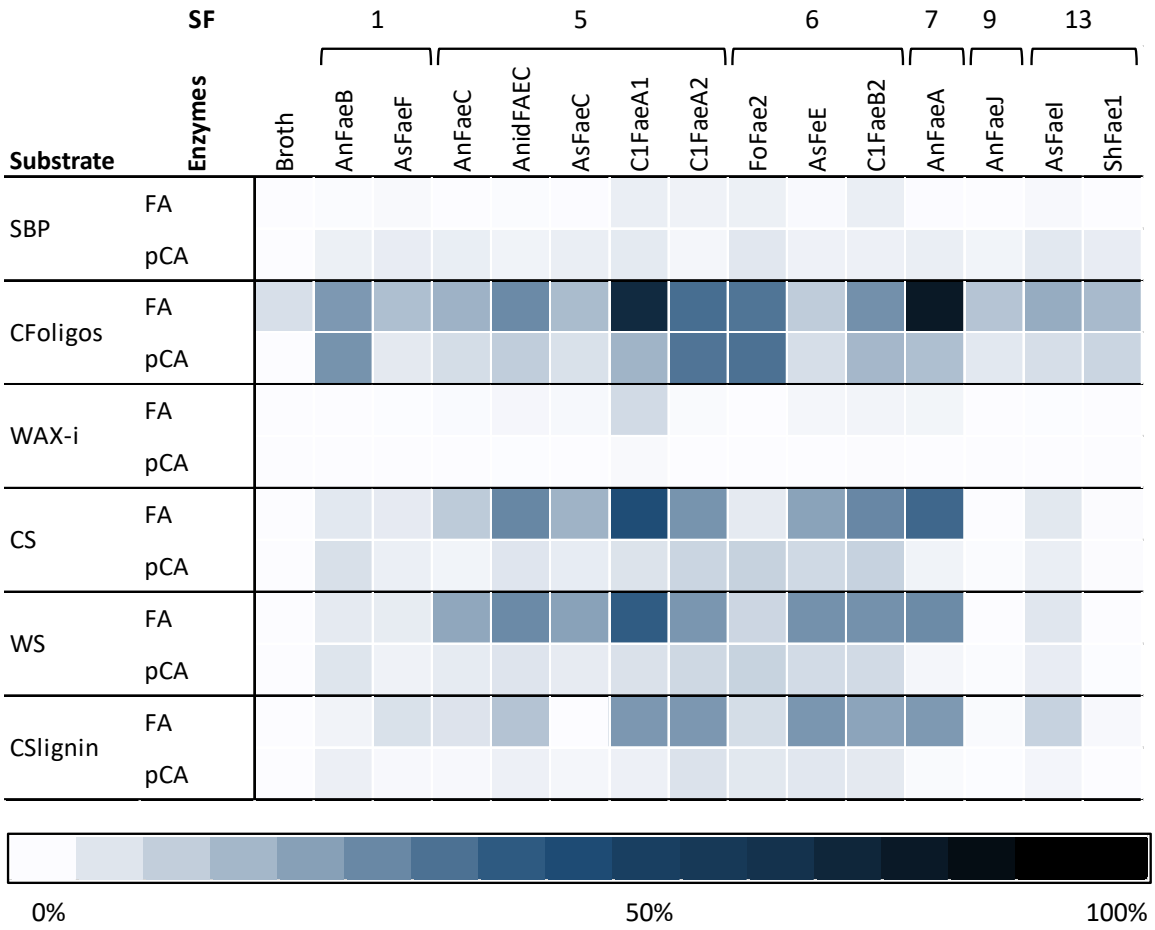


Figure 4-4: Heatmap for the release of FA and pCA from the natural substrates tested, the results are shown as percentages of bound content (0 – 100%) as measured by UHPLC-UV. CS: corn stover; WS: wheat straw.

Additionally the release of various di- and triFAs from the natural substrates were quantified. The release of the polymeric compounds from CFoligo, corn stover, wheat straw, and CSlignin can be seen in Figure 4-5 (for numerical values and standard deviations see Table A3 in Appendix).

For the incubations with the four aforementioned natural substrates, the highest release of diFAs was observed for the enzymes categorised as SF5 and 7. It was furthermore seen that the diFAs 8-8'-(furan)-diFA, m/z = $401_{9.75\ min}$, and 5-5'-diFA were the main diFAs released due to FAE activity. Interestingly, diFAs were even released from CSlignin.

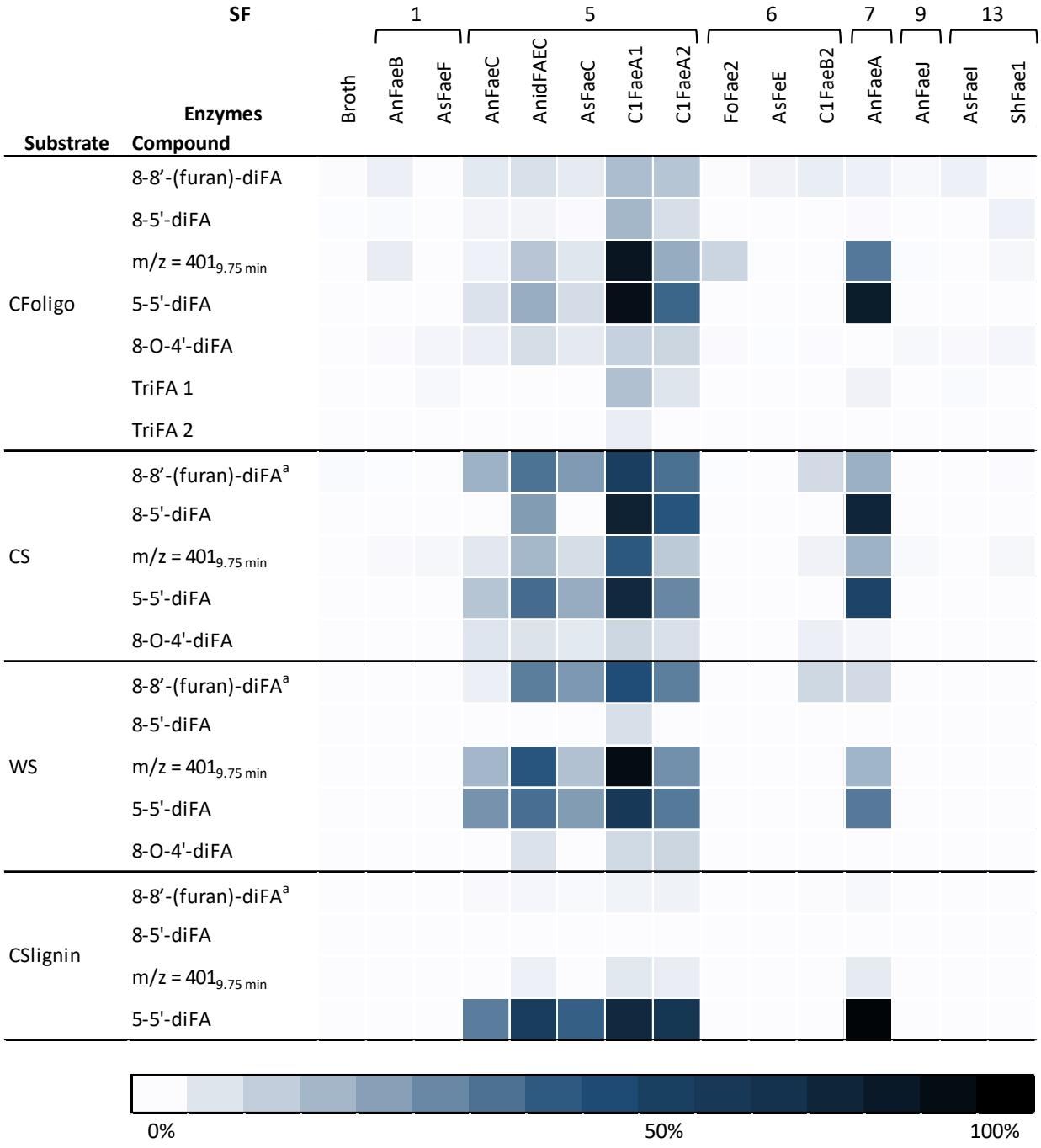


Figure 4-5: Heatmap showing for the release of diFA, triFA and m/z = 401 from CFoligo, corn stover (CS), wheat straw (WS), and CSlignin as percentages of the ester-bound content (0 - 100%) as measured by UHPLC-MS. [a]m/z = 389, corresponding to a reduced form of 8-8'-(furan)-diFA.

The incubation with SBP and WAX-i released either no or very low amount of the polymeric compounds (Figure 4-6, for numerical values and standard deviations see Table A3 in Appendix).

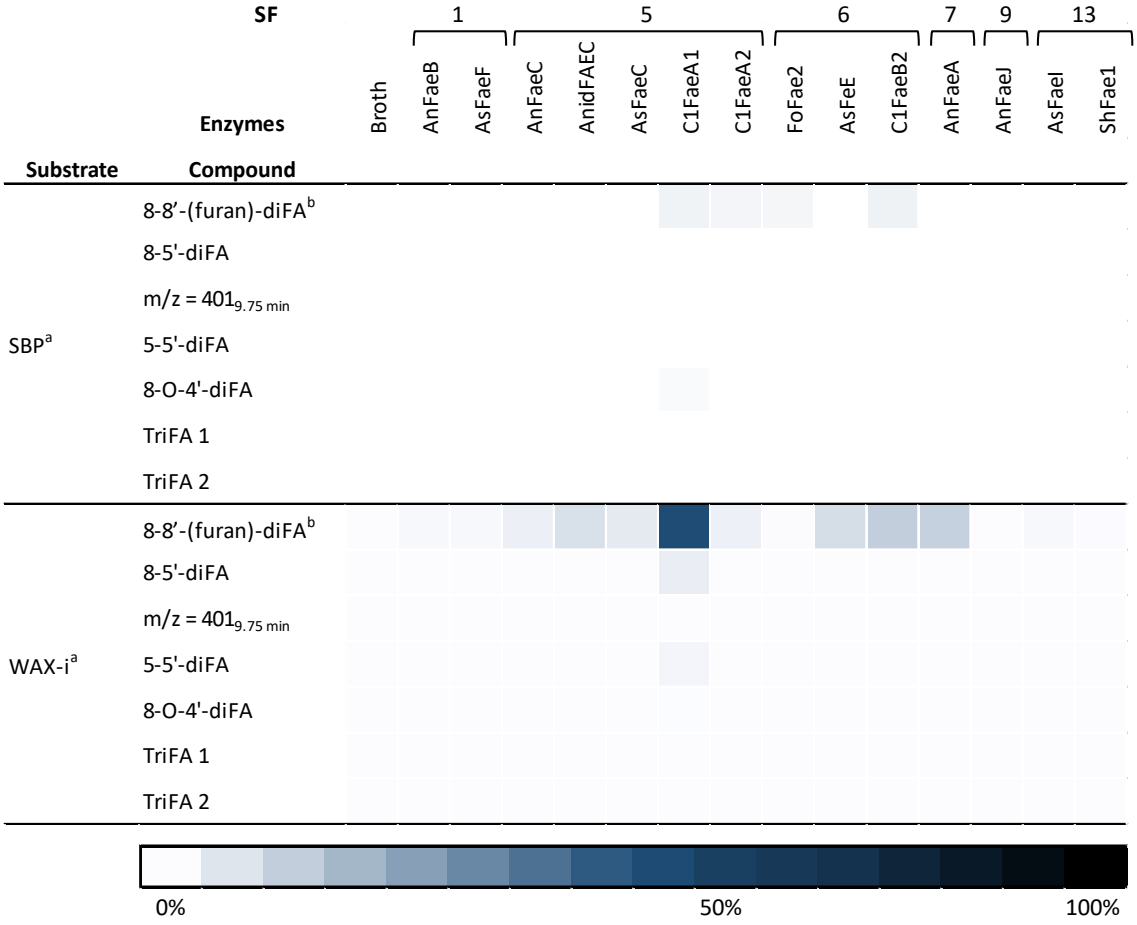


Figure 4-6: Heatmap presenting the release of diFA, triFAs and m/z = 401 from SBP and WAX-i as percentages of bound content (0 - 100%) as measured by UHPLC-MS. [a]Quantified by UV. [b]m/z = 389, corresponding to a reduced form of 8-8'-(furan)-diFA.

The absolute values for the release of m/z = 401$_{10.66\ min}$ are shown in Figure 4-7 (for numerical values see Table A4 in Appendix) as the compound was unstable during saponification for which reason the total amount of the compound could not be determined. Apart from CFoligo, the compound was released in similar amounts from corn stover, wheat straw, and CSlignin by the enzymes categorised as SF5 and 7 enzymes. Also the SF6 enzyme C1FaeB2 showed activity towards this compound for the four aforementioned substrates.

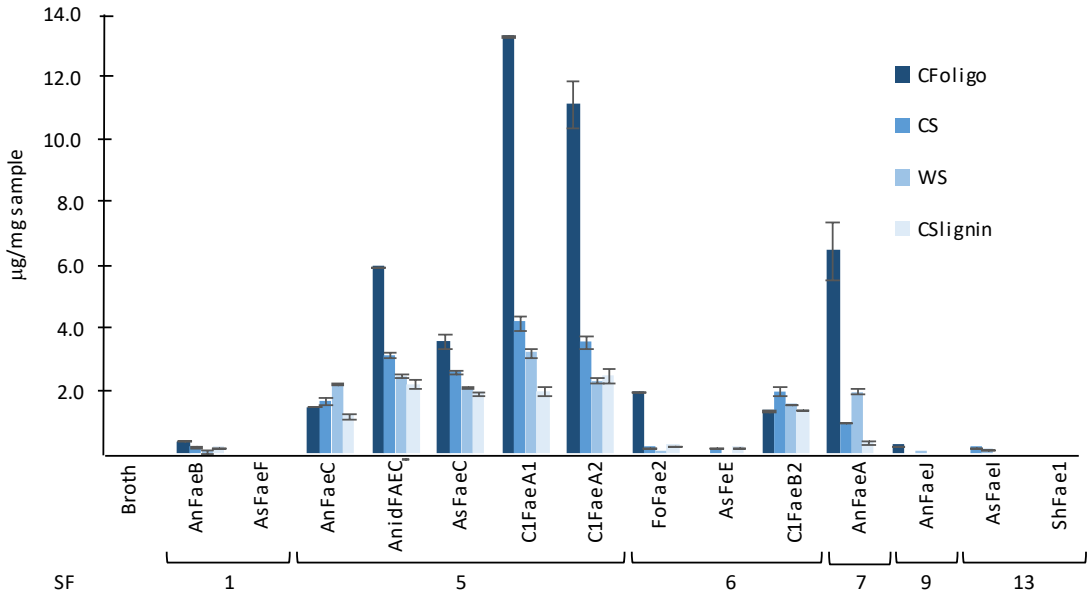


Figure 4-7: Release of m/z = $401_{10.66\ min}$ as absolute value in µg/mg for the natural substrates CFoligo, corn stover (CS), wheat straw (WS), and CSlignin as measured by UHPLC-MS.

4.5 Discussion of the observed FAE activities

In this study, the hydrolytic activity of 14 FAEs towards the model substrates MF and MpC have been determined by RP-UHPLC-UV. Based on the results the SF1 enzymes were moderately active towards the hydrolysis of MF (2 and 19 h incubation), but more so towards MpC, where the degradation of the substrate was almost complete after 19 h incubation (Figure 4-1). These results are in line with previously reported data, where SF1 is indicated to prefer smaller substrates, also fitting with the ABCD classification of these FAEs as type B (Table 4-1 and 1-1).[27,66,180]

Among the all 14 tested FAEs, the SF5 enzymes were the most active, where generally 50% hydrolysis was achieved after only 2 h for both model substrates, and full degradation of MF was observed after only 2 h for C1FaeA1 (Figure 4-1). This corresponds well with the previous classification as type A and C or D enzymes (Table 4-1), where type A prefers methoxy substituents and type C and D show general activity (Table 1-1). However, the results for enzyme AnidFAEC (SF5) did not follow the same trend, and was completely inactive towards the tested model substrates (Figure 4-1), which is not in line with the previously published data[192] as well as the enzyme being a type C FAE according to the ABCD classification (Table 4-1). An explanation for the contradictory results can be substrate inhibition, as this study was done at much higher substrate concentration than previous reports (approximately 20 times higher).[192] This assumption was corroborated by the de Vires laboratory, as substrate

inhibition have also previously been observed for this enzyme (unpublished results). Nevertheless, the enzymes is still considered a true FAE as it released FA, pCA, and diFAs from the natural substrates.

The FAE enzymes of SF6 demonstrated activity towards both tested model substrates, although for FoFae2 the hydrolytic activity decreased after 2 h (Figure 4-1). AnFaeA of SF7 was inactive towards degradation of MpC, but it was highly active towards the hydrolysis of MF, where 75% degradation was observed after 19 h (Figure 4-1). This could be indicative of a preference for more bulky substrates even though according to the previous type A classification of this enzyme then smaller substrates should also be hydrolysed (Table 4-1 and 1-1).[27] The one tested FAE of SF9 was showed no hydrolytic activity towards both MF and MpC, while for the FAEs of SF13 the only active enzyme was AsFaeI (Figure 4-1). Overall these results are in accordance with the previously published data.[66,179]

In summary, the FAEs of SF1 and SF6 were most active towards MpC, but still also active towards MF; SF5 FAEs were active towards both substrates; SF7s were not active towards MpC but were towards MF; SF9s were inactive towards both model substrates; and SF13 FAEs were slightly active towards both MF and MpC.

The identity of the quantified commonly described diFAs was confirmed by MS^2 fragmentation and the order of elution, as these are in accordance with the previous published data[31] as well as results from the group at the Food Chemistry Department (unpublished data) (Figure 4-2, Table 4-5, and Figure A1-A8). In addition, the two compounds with m/z = 401 is proposed to be regioisomers of FA and 5-hydroxyferulic acid dimer based on MS characterisation and elution pattern (Figure 4-2, Table 4-5, Figure A5 and A7).

SBP and WAX-i were generally found to be poor substrates for the tested FAEs, as no or very low release of the quantified compounds were observed (Figure 4-4 and 4-6). It has previously been described that FAE activity on insoluble substrates is dependent on the presence of a carbohydrate binding module,[199] and therefore pre-treatment with e.g. xylanases are needed before activity is observed.[179] However, in this study the substrates corn stover and wheat straw were also insoluble in the sodium acetate buffer used, but several of the FAEs tested still exhibited pronounced activity in line with the activity observed towards the fully soluble CFoligo substrate. Hence, this study shows that substrate solubility can not completely explain the lower FAE activity towards SBP and WAX-i (Figure 4-4 and 4-6).

The SF1 enzymes were less active towards the natural substrates CFoligo, corn stover, wheat straw, and CSlignin than towards the model substrates. They were able to release small amounts of FA and pCA but no hydrolytic activity was observed towards diFAs and triFAs (Figure 4-4 and 4-5). Compared to the FAEs of SF1, the tested FAEs of SF13 enzymes displayed similar behaviour not dependent on the substrate. This is in agreement with their previous classification as type B FAEs (Table 4-1), however their further subdivision as SF1 and 13 seems justified based their different catalytic behaviour towards especially the model substrate MpC.

The tested FAE belonging to SF9 was inactive towards all insoluble natural substrates like the FAEs from SF1 and SF13. The soluble CFoligo was the only substrate, from which relatively low release of FA was observed, but no pCA was released. Hence, compared to SF1 and 13 enzymes distinctive activity of the SF9 enzyme was observed. An additional explanation for the inactivity of the SF1 and SF9 enzymes can be the sequence similarity to tannases.[27]

The FAEs of SF5, 6, and 7 were all able to release FA from CFoligo, corn stover, wheat straw, and CSlignin, where the SF5 FAEs exhibited most pronounced activity (Figure 4-4). Furthermore, was pCA released from the four aforementioned natural substrates by the enzymes of SF5 and 6, while the one SF7 enzyme showed distinctive behaviour, as it only released pCA from the soluble CFoligo.

In contrast to all other tested SFs SF5 and 7 enzymes were able to release various diFAs and triFAs. Hence, the tested FAEs belonging to SF6 were hardly able to release diFAs and triFAs, except for the incubation of especially corn stover with C1FaeB2, which showed low activity towards some of the diFAs (Figure 4-5 and 4-7). The limited ability of C1FaeB2 for the release of diFA are in line previous results.[179] However, the general inability of the SF6 enzymes to hydrolyse polymeric ferulic acids are in accordance with the previous classifications as type B enzymes (Table 4-1).[27,63] This furthermore shows, that the enzymes of SF5 and 6 are clearly distinct even though they belong to the same CE1 family in CAZy.

Interestingly the two SF6 enzymes AsFaeE and FoFae2 exhibited opposite substrate specificity. When these enzymes were incubated with the model substrates, and the natural substrates corn stover, wheat straw, and CSlignin AsFaeE was more active towards the release of FA and pCA than FoFae2, but the opposite was seen when the incubated substrate was CFoligo.

The new classification of the FAEs into 13 SFs shows its strength especially in the grouping of the SF5 enzymes, which showed similar ability for the release of FA, pCA, diFAs, and triFAs from

both model and natural substrates (Figure 4-4, 4-5, 4-6). However, according to the ABCD classification these enzymes would be classified as several different types (A, C, C/D) (Table 4-1).[27,180] Hence, based on these results the co-classification as SF5 enzymes seems realistic.

The diFAs were released to different degrees by all the enzymes categorised as SF5 and 7, although triFAs were only released by the enzymes C1FaeA1, C1FaeA2, and C1FaeB2 (Figure 4-5). Curiously, these were also the only tested enzymes, which were fully purified, and triFA-activity might therefore be related to the level of purification. Thus, some of the other enzymes might also show activity towards triFAs if purified to a similar degree.

Beyond comparison of the catalytic behaviour of the FAEs, a general trend for preferential FA release compared to pCA release from the natural substrates were observed especially for SF5, 6, and 7 (Figure 4-4). This was even the case when rather high amounts of ester-bound pCA was present (98.25 and 28.39 μg/mg sample for CSlignin and corn stover, respectively). A possible explanation for this phenomenon for CSlignin at least could be that pCA in lignin is not well accessible for the FAEs.[32,200]

Looking specifically at the diFAs released by FAE activity, a preference in the type of structure was seen for both SF5 and 7. Generally, 5-5'-diFA was the most released from CFoligo, corn stover and wheat straw (up to 91% for CFoligo) followed by 8-8'-(furan)-diFA (up to 61% for corn stover), and m/z = $401_{9.75\ min}$ (up to 92% and 87% for wheat straw and CFoligo, respectively) (Figure 4-5, Table A3). From CSlignin the 5-5'diFA was also the most released diFA (up to 91% for CSlignin), while low release of 8-8'-(furan)-diFA was observed even though it was the most abundant diFA of the substrate (45.93 μg/mg sample), and the reason might again be ascribed to the complexity of lignin.

Additionally was the 8-5'-diFA was released to a higher extend from corn stover than from CFoligo even though a higher amount of the diFA is present in the latter (13.83 compared to 56.63 μg/mg sample for corn stover and CFoligo, respectively) (Figure 4-5, Table A3). This could indicate that 8-5'-diFAs bound to GAX in CFoligo were less readily accessible, than 8-5'diFAs present in corn stover, which represented a more generaæ composition of lignocellulosic material.

Generally the FAEs were not able to release the diFA 8-O-4' to any great extend independent of the substrate even when high amounts were present (153.83 μg/mg sample for CFoligo) (Figure 4-5, Table A3). Thus, indicating that this structure for some reason was less accessible

than the other diFAs, which might be the result of either the catalytic site structure of the FAEs or of the substrate inaccessibility.

4.6 Concluding remarks

In this study the ability of 14 different FAEs to hydrolyse both monomeric and polymeric hydroxycinnamic esters from different substrates have been investigated. The results were analysed by RP-UHPLC-UV-ESI-MS providing more accurate compared to the commonly used spectrophotometric method.

It can be concluded that the FAE activity of an enzyme can be substrate and product specific, as well as the solubility of the substrate is not necessarily a determining factor for activity. Additionally several of the enzymes were proven active even towards a purified lignin isolate, showing that lignin contains several ester-linked diFAs cleavable by FAEs.

The strength of the 13 SF classification of the FAEs is seen in this study, where subtle differences in reactivity is recognised in the 13 SF classification. This is especially seen by the grouping of the SF5 enzymes, which show similar reactivity but according to the ABCD classification, they would be classified as several different types.

Based on this study the SF5 and SF7 enzymes showed the broadest FAE activity both towards monomeric and polymeric hydroxycinnamic esters over a variety of natural substrates. Hence, these enzymes or other enzymes of these families would be interesting for further studies, as well as being interesting candidates for industrial purposes. Nevertheless, since only one SF7 enzyme has been tested, other SF7 enzymes would be interesting to test and determine if the reactivity observed is a general trend for this SF.

4.7 Experimental

General considerations

UV measurements were performed using TECAN – Infinite F500 with 99 well microplates. Centrifugations were done in a Hermle Z383K centrifuge or Hermle Z233 MK-2. Thermo mixers used were Eppendorf® Thermomixer Comfort. Boiler utilised was a Julabo TW8. All solvents used were of analytical HPLC grade.

Materials

Substrates: Methyl ferulate and methyl *p*-coumarate were obtained from Apin Chemicals Ltd. and Carbosynth, respectively. Ferulic acid and *p*-coumaric acid were obtained from Fluka™ and SigmaAldrich, respectively. Sugar beet pectin (SBP, Pectin Betapec RU301, Herbstreith & Fox KG), wheat arabinoxylan insoluble (WAX-i, Megazyme), corn fiber oligomers (CFoligo),[182] corn stover (CS) and wheat straw (WS) were kindly provided by CNC (Milsbeek, The Netherlands).

Enzymes: The expression, purification and characterization of C1FaeA1 (JF826027, subfamily 5 (SF5; Dilokpimol et al., 2016)), C1FaeA2 (JF826028, SF5), and C1FaeB2 (JF826029, SF6) from *Chrysosporium lucknowense* C1 have been described previously.[179] The expression and production of the FAEs from *Aspergillus niger*: AnFaeB (CAC83933.1, SF1), AnFaeC (XP_001395336.1, SF5), AnFaeA (CAA70510.1, SF7), and AnFaeJ (An15g05280, SF9), from *Aspergillus sydowii*: AsFaeF (Aspsy1_293049, SF1), AsFaeC (Aspsy1_154482, SF5), AsFaeE (Aspsy1_1158585, SF6), and AsFaeI (Aspsy1_160668, SF13), from *Aspergillus nidulans*: AnidFAEC (AN5267, SF5), from *Fusarium oxysporum*: FoFae2 (Fusox1_8990, SF6), and *Stereum hirsutum*: ShFae1 (Stehi1_73641, SF13) has been described previously.[66]

Enzyme filtering and measurement of protein content

Approximately 7.5 mL of the culture filtrate was concentrated to approx. 2 mL using ultrafiltration (Amicon Ultra, molecular mass cut-off of 3 kDa, Merck Millipore). The filtrate was removed, and the concentrate diluted with sodium acetate buffer (50 mM, pH 5.8) to about 7 mL and concentrated again using ultrafiltration. This washing procedure was performed twice. The protein content of the concentrated samples were measured using the BCA Protein Assay Kit – Reducing agent compatibility (Thermo Scientific) and bovine serum albumin (BSA) as calibration. The protein content of the broth from *P. pastoris* without FAE insertion (negative control, coded as 'broth') was measured with BCA Protein Assay Kit – Reducing agent

compatibility (Thermo Scientific). After protein content determination, all enzymes were diluted to a final protein concentration of 1 mg/mL in a sodium acetate buffer (50 mM, pH 5.8).

Incubation of FAEs with model substrates

Activity assays towards the model substrates MF and MpC was performed in 200 µL reaction mixtures. Therefore, 2 mM substrate was dissolved in a sodium acetate buffer (50 mM, pH 5.8) (175 µL) and 25 µg of FAEs-containing filtrates (1 mg/mL) or the broth (1 mg/mL) was added. In contrast, 1 µg of the purified FAEs from *Chrysosporium lucknowense* C1 (C1FaeA1, C1FaeA2, and C1FaeB2) was added. The reaction mixtures were incubated at 37 °C for 2 h and 19 h. All reactions were performed in duplicate. At 2h and 19h, 50 µL was taken from the reaction mixture and diluted 10 times with MilliQ. All reactions were stopped by incubating the samples at 99 °C for 2 min, and appropriately diluted before UHPLC-UV analysis.

Determination of free and bound hydroxycinnamic acids

The maximum release (free and bound content) of hydroxycinnamic acids in the natural substrates were estimated based on amounts released in 0.25 M KOH. Hereto, in triplicate, 5 mg of substrate (or 1 mg for CSlignin), 500 µL of acetate buffer (or 100 µL for CSlignin; 50 mM, pH 5.8) and 500 µL (or 100 µL for CSlignin) of 0.5 M KOH was mixed. The free amount of hydroxycinnamic acids in the natural substrates were estimated, in duplicates, by mixing 5 mg of substrate (or 1 mg for CSlignin), and 500 µL of acetate buffer (or 100 µL for CSlignin; 50 mM, pH 5.8). All samples were incubated for 19 h in the dark at 37 °C, head-over-tail. After incubation saponified samples were acidified (pH < 3) by addition of 6 M HCl. All samples were centrifuged (12000 *g*, 10 min, 4°C), and appropriately diluted before UHPLC-UV-ESI-MS analysis.

Enzyme activity assays with natural substrates

Enzyme activity assays towards natural substrates were performed for all 14 FAEs and the broth. Insoluble substrates SBP, CS, and WS were weighed per incubation. CFoligo was dissolved in acetate buffer (50 mM, pH 5.8) to a stock solution of 10 mg/mL. CSlignin was dissolved in EtOH/$CHCl_3$ 50:50 to 1 mg/mL and after distribution of sample-material the solvent was evaporated. In comparison, CSlignin was used at a lower concentration, due to substrate limitation. The reactions were performed in the presence of 5 mg of substrate (or 1 mg of CSlignin) in 500 µL (or 100 µL for CSlignin) of a sodium acetate buffer (50 mM, pH 5.8), enzyme-protein loading was 25 µg for FAE-containing filtrates (1 mg/mL), and 2.5 µg enzyme

for purified FAEs from *Chrysosporium lucknowense* C1 (C1FaeA1, C1FaeA2, and C1FaeB2). All reactions were done in duplicates. The samples were incubated for 19 h in the dark at 37 °C, head-over-tail. Enzymes were inactivated at 96 °C for 5 min. All samples were centrifuged (12000 *g*, 10 min, 4°C), and appropriately diluted before UPLC-UV-ESI-MS analysis. The amount of hydroxycinnamic acids removed were corrected for the free amount of hydroxycinnamic acids present.

RP-UHPLC-UV-ESI-MS/MS analysis

Analysis of the model substrates MF and MpC incubated with FAEs. All samples were subjected to an Accela reversed phase ultra-high performance liquid chromatography (RP-UHPLC) system coupled to electron spray ionization (ESI) mass spectrometry (Thermo Scientific). Samples (5 µL) were separated using an Acquity UPLC BEH C18 column (150 x 2.1 mm, particle size 1.7 µm) (Waters). The flow rate was 400 µL/min at 45 °C. The binary mobile phase consisted of (A) water + 0.1% formic acid and (B) acetonitrile + 0.1% formic acid. The elution profile was as follows: Isocratic on 5% B; 0.0 – 1.5 min, B linearly from 5% to 60%; 1.5 – 20.0 min, B linearly from 60% to 100% B; 20.0 – 20.1 min, isocratic on 100% B; 20.1 - 25.0 min, B linearly from 100% to 5% B; 25.0 – 26.0 min, isocratic on 5% B; 26.0 – 31.0 min. The flow rate was 0.400 mL/min. UV spectra were recorded at 320 nm. The decrease in substrate concentration representing activity was determined from a standard curve of the substrates (0.625 - 50 µg/mL). No MS data acquired for the incubation of the model substrates with the FAEs. Data were processed using Xcalibur 2.2 (Thermo Scientific).

Analysis of the natural substrates for maximum and free content as well as for incubations with FAEs. Samples were analysed by using RP-UHPLC-UV-ESI-MS as described above with a modified elution profile. The binary mobile phase consisted of (A) water + 0.1% formic acid and (B) acetonitrile + 0.1% formic acid. The elution profile was as follows: The first two minutes isocratic on 5% B; 0.0 - 2.0 min, B linearly from 5% to 40%; 2.0 – 13.0 min, B linearly from 40% to 100% B; 13.0 – 13.1 min, isocratic on 100% B; 13.1 - 18.0 min, B linearly from 100% to 5% B; 18.0 – 18.1 min, isocratic on 5% B; 18.1 – 23.0 min. The flow rate was 0.400 mL/min. UV spectra were recorded at 280 and 320 nm. Samples were further analysed using an LTQ-Velos mass spectrometer (Thermo Scientific) equipped with an ESI-MS. Data was collected over a m/z range of 120 - 1500 in negative mode. MS settings were optimised with LTQ-tune. Nitrogen was used as sheath gas and auxiliary gas. Data dependent MS^2 analysis was performed using collision-induced dissociation with a normalized collision energy of 35%. The ion transfer tube

temperature was 300 °C, source heater temperature was 250 °C and the source voltage was 3.5 kV. Data were processed using Xcalibur 2.2 (Thermo Scientific). Due to linear dependency of the MS signals to the concentration for FA and pCA, the quantification of diFA and triFA was done by MS for all natural substrates except for SBP and WAX-i where the UV signal was used, as low amounts of the quantified compounds where released. The contents of diFA and triFA were calculated relative to FA and mass corrected.

5. Regioselective glycosylations

In this chapter a brief overview of previously published regioselective glycosylation will be given, followed by a more specific short review over previous investigations of molybdenum complexations. The last sections will present the effort invested in pursuing regioselective glycosylations with molybdenum and concluding remarks.

5.1 Introduction to regioselective glycosylations

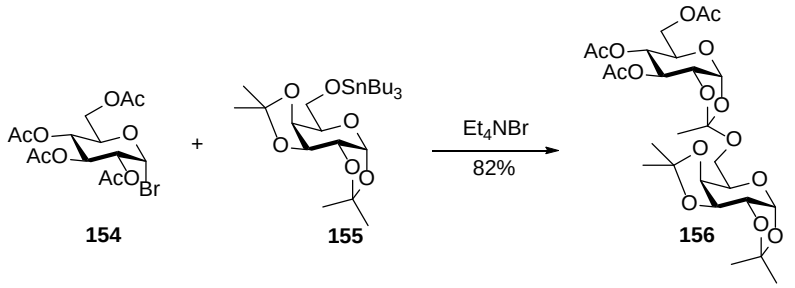
Synthetic glycosylation chemistry is especially known for the necessity of tedious protection and deprotection reactions often leaving only one hydroxyl group of the glycosyl acceptor unprotected, to ensure regio- and stereoselective glycosylations as also observed in the previous chapters.[201] The need for protection groups can be exemplified by the fact that 1056 trisaccharides can be assembled from only 3 different monosaccharides.[202] Hence many elaborate strategies for protections have been developed through the years, but the many protection and deprotections add multiple steps to the synthetic approach. Consequently, there have been a great interest in reducing the number of protection groups thus shorten the overall synthesis. A solution is regioselective manipulation of partial and completely unprotected carbohydrates.

The inherent difference in reactivity of hydroxyls in carbohydrates is a key factor in regioselective manipulations. Generally, primary alcohols are more reactive and can therefore

undergo selective manipulations over secondary alcohols. Additionally, are equatorial secondary hydroxyls are more reactive than axial hydroxyls. Furthermore, it is known that transition metals can form complexes with carbohydrates.[203,204] These complexations can be assumed to influence the nucleophilicity of the involved hydroxyls in a way, which is depending on the metal and the geometry of the intermediate complex. Hence, complexation increases the inherent discrimination of reactivity. This has been used for selective placement of protection groups leading to selective alkylations, esterifications, phosphorylation, and sulfonylation.[205–209] Building on these experiences the field of regioselective glycosylations has been developed.

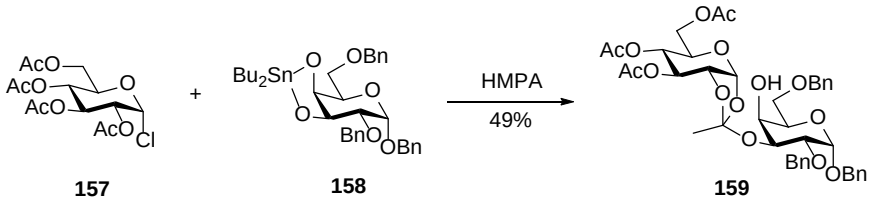
5.1.1 Tin-mediated glycosylations

The first glycosylation applying an organotin(IV) reagent was achieved by Ogawa and Matsui in 1976.[210] The glycosylation was performed using acetobromoglucose (**154**) and the galactose derived trialkyltin reagent **155** resulting in orthoester product **156** shown in Scheme 5-1.



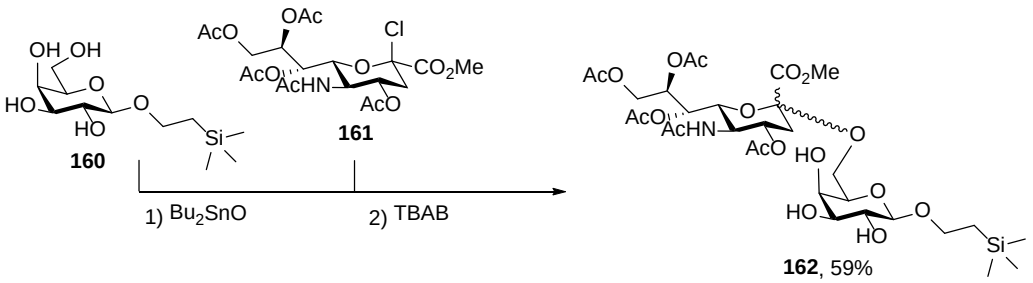
Scheme 5-1: Glycosylation by Ogawa and Matsu employing a trialkyltin reagent.[210]

However, this glycosylation cannot truly be called regioselective, as only the 6-position of the acceptor could be glycosylated. The first regioselective glycosylation was achieved a few years later in 1979 by Augé and Veyiéres[211] inspired by the work of Ogawa and Matsu. The 1979 example employed dialkylstannylene acetals to selectively achieve the 3-linked orthoester **159** from the acetochloroglucose (**157**) (Scheme 5-2). However, the same selectivity was observed in the absence of the stannylene activation, and the reaction was therefore not tin-mediated.



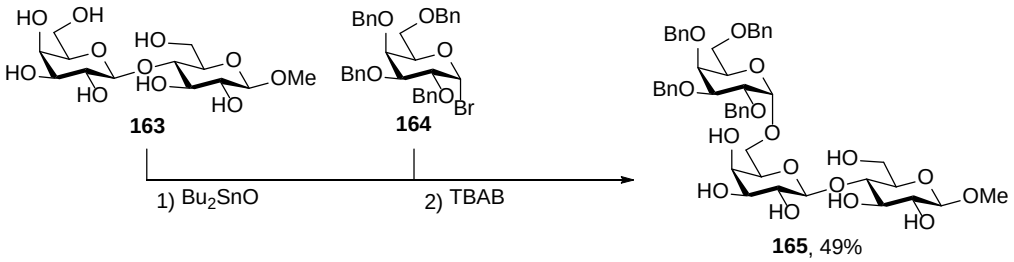
Scheme 5-2: Glycosylation by application of stannylene acetals.[211]

Years later Hasegawa and co-workers[212] published stannylene acetal formation of unprotected 2-(trimethylsilyl)-ethyl β-D-galactopyranoside (**160**), which was utilised for glycosylation with acetochlorosialic acid **161** for which orthoester formation was impossible. The glycosylation was regio- but not stereoselective and resulted in an anomeric mixture of the 1,6-linked product **162** in 59 % yield (Scheme 5-3).



Scheme 5-3: Stannylene mediated glycosylation as reported by Hasegawa and co-workers.[212]

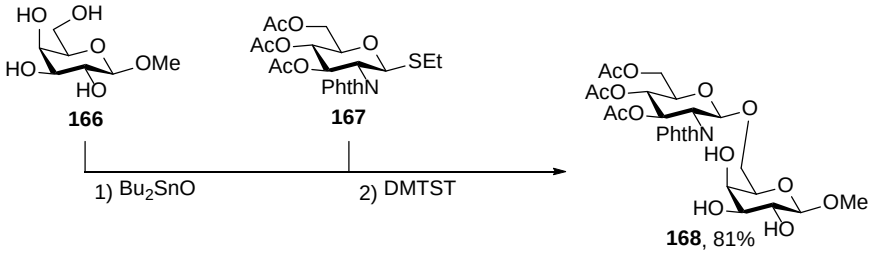
However, a regio- and stereoselective glycosylation was obtained by Martin-Lomas and co-workers.[213] Here methyl β-lactoside (**163**) was selectively glycosylated at the 6-position with 2,3,4,6-tetra-*O*-benzyl-α-D-galactopyranosyl bromide (**164**), yielding the trisaccharide **165** in good yield under similar conditions (Scheme 5-4).



Scheme 5-4: Tin-mediated glycosylation of methyl β-lactoside in a regioselective manner.[213]

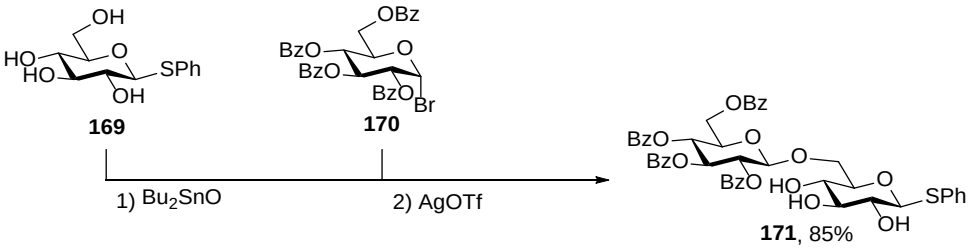
Oscarson and co-workers expanded the stannylene method in a publication from 1995.[214] Here the stannylene acetals were employed to achieve formation of 1,6-linkages between different protected thioglycosides and glycosyl bromides as donors and the fully deprotected methyl β-D-galactopyranoside (**166**) as the acceptor. The setup achieved moderate to good yields. As the best example ethyl 3,4,6-tri-*O*-acetyl-2-deoxy-2-phthalimido-1-thio-β-D-glycopyranoside (**167**) was employed as the donor. Glycosylation of the stannylene derived galactopyranoside provided the 1,6-linked product **168** in 81% yield (Scheme 5-5). The influence of the tin-reagent on the regioselectivity was confirmed by a control experiment without addition of the

tin reagent, where a mixture of tri- and tetra-saccharides with no apparent selectivity were observed.



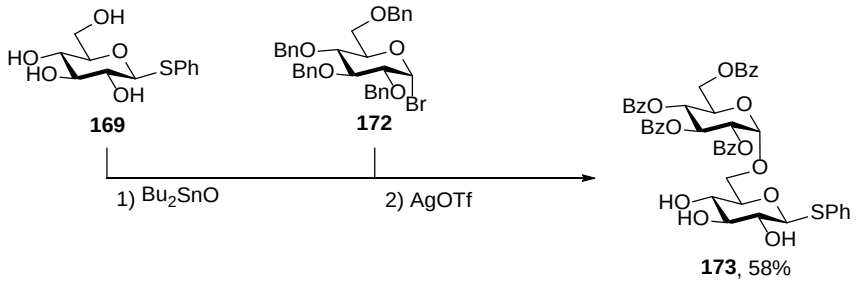
Scheme 5-5: Example of the Oscarson contribution to the field.[214]

Many have contributed to the field of tin-mediated glycosylations, and more recently, the Madsen group provided a further elaboration of the field. Maggi and Madsen[215] developed a protocol for 1,6-glycosylation of different thiophenyl glycoside acceptors with glycosyl bromides of glucose, galactose, mannose, and glucosamine in up to 85% yield (Scheme 5-6). For this protocol, dibutyltin oxide was used as the mediator, silver triflate as the promoter and 4Å MS acted as a desiccant and a weak acid scavenger.



Scheme 5-6: The best example of tin-mediated glycosylation as by Maggi and Madsen.[215]

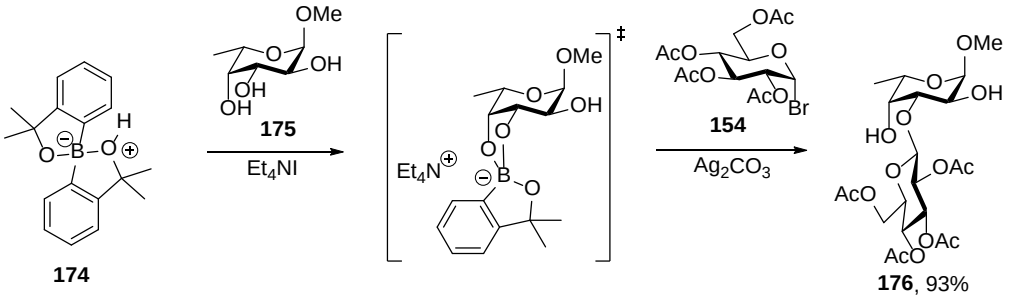
In 2016 Niedbal and Madsen[216] explored stannylene-mediated glycosylations obtaining a number of 1,2-*cis*-linked disaccharides. This was done by employing perbenzylated hexopyranosyl bromides as donors and TBAB as the promotor, providing the 1,6-linked product in moderate yield as exemplified in Scheme 5-7. Generally, the reactions were limited by the solubility of the acceptors in the reaction solvent.



Scheme 5-7: α-1,6-linkage as formed by regioselective glycosylation mediated by stannylene.[216]

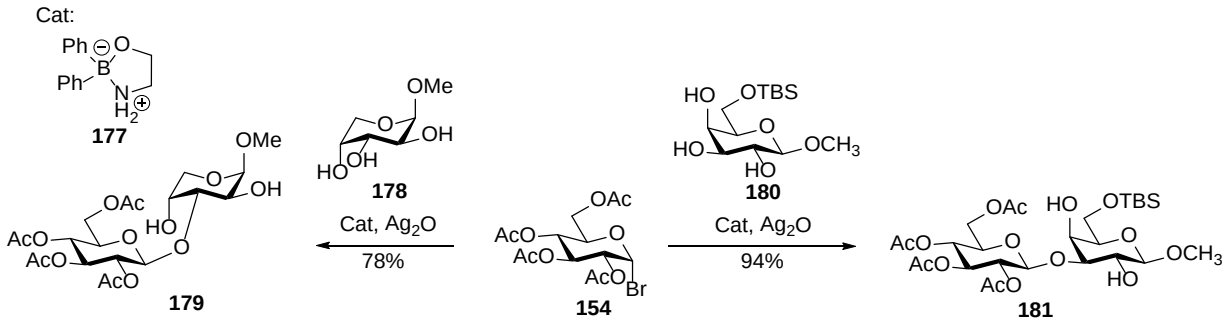
5.1.2 Boron-mediated glycosylations

Boronic acid complexes of carbohydrates have also been utilised for regioselective glycosylations. The first example hereof was reported in 1999 by Oshima and Aoyama.[217] Cyclized 2-(2-hydroxy-2-propyl) phenylboronic acid (**174**) was used as the promotor to achieve glycosylation between acetobromoglucose (**154**) and several unprotected and partly protected hexopyranoses in the presence of tetraethylammonium iodide and silver carbonate. Regioselectivity was achieved for 1,2-diols and 1,3-diols involving a primary hydroxyl group through a boron-acetal, activating the least hindered B-O moiety leading to glycosylation as exemplified in Scheme 5-8.



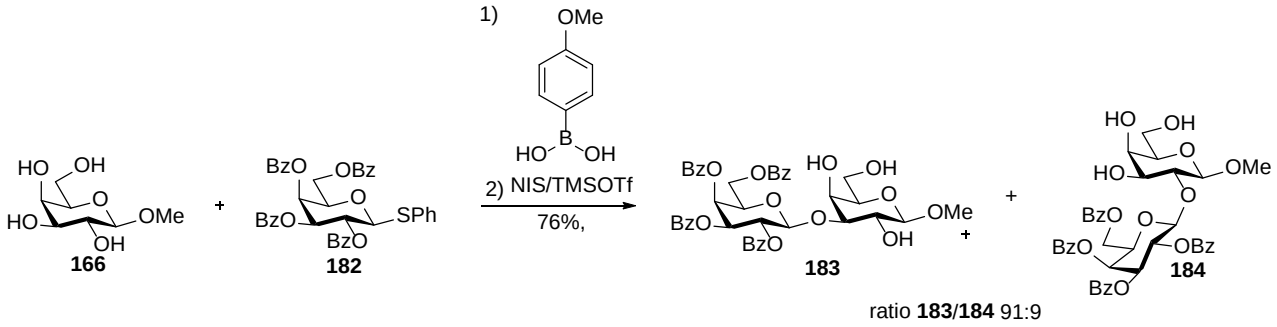
Scheme 5-8: Regioselective glycosylation as by Oshima and Aoyama.[217]

The same selectivity for glycosylations through activation of the equatorial hydroxyl group of a 1,2-*cis*-diol motifs were also demonstrated by Taylor and co-workers in 2011.[218] The Koenigs-Knorr glycosylations were promoted by a diphenylborinic acid (**177**) providing good yields for a range of different halide donors and acceptors derived from galactose, mannose, fucose, and arabinose as illustrated in Scheme 5-9. In the case of a primary hydroxyl in the acceptor, the 6-position was protected to avoid possible binding of the boron-catalyst to the *O*-4 and *O*-6.



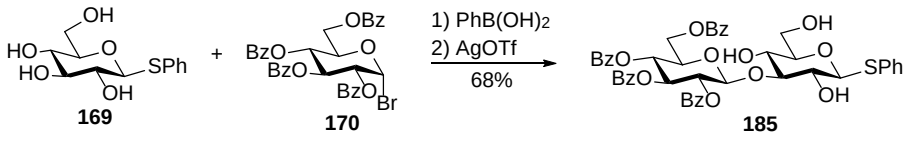
Scheme 5-9: Examples of borinic acid promoted glycosylation as performed by Taylor and co-workers.[218]

The examples above display the ability of boron reagents to activate a specific position of the acceptor which led to regioselective glycosylations. However, arylboronic acids can be used in an alternative procedure by transiently masking the corresponding hydroxyl groups and in this way making attack of the most reactive hydroxyl possible. One of the first that explored this was Kaji and co-workers.[219] Using glycosyl bromides and phenyl thioglycosides as donors glycosylations with fully unprotected hexopyranosides were achieved in 66-78% yield with great regioselectivity for the 1,3-linked disaccharide. As an example hereof see Scheme 5-10.



Scheme 5-10: Using a boronic acid for transient masking.[219]

The same type of *in situ* masking was explored by Fenger and Madsen.[220] Several thioglycoside acceptors were treated with phenylboronic acid installing the temporary boronic ester and subsequently submitted to Koenigs-Knorr glycosylations with perbenzoylated gluco- or galactopyranosyl bromides providing the 1,3-linked disaccharides. The model reaction is shown in Scheme 5-11. This procedure was furthermore utilised to obtain different 1,3-linked trisaccharides in a one-pot style in up to 66% yield.

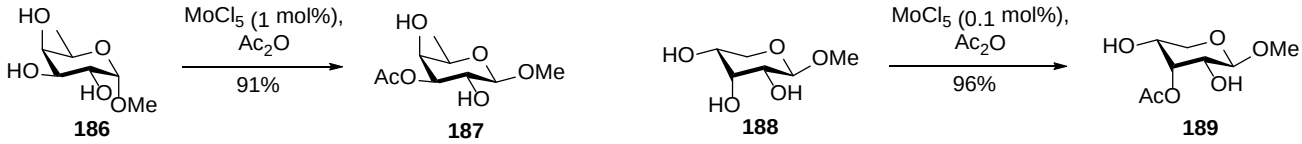


Scheme 5-11: Model reaction of the Fenger and Madsen *in situ* masking with boronic acid.[220]

In summary, literature provides several examples of more or less successful regioselective glycosylations by different means. Only a few cases demonstrate how completely unprotected carbohydrates can be used for glycosylations, and despite years of research a universal method has not yet been developed. Furthermore, there is still room for improvement and novel procedures would be highly desirable.

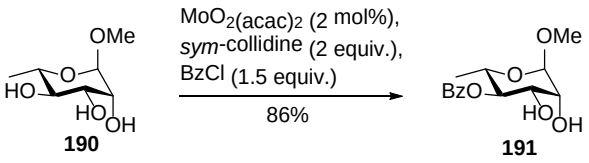
5.2 Possibility of molybdenum-mediated glycosylations

Different molybdenum complexes have been shown to mediate regioselective protection group manipulations by E. Evtushenko.[206,207] The first example demonstrated the ability of $MoCl_5$ to mediate regioselective monoacetylation of several carbohydrates containing 1,2-*cis* vicinal hydroxyls in good yields. The best results were obtained using methyl α-D-fucopyranoside (**186**) and methyl β-D-ribopyranoside (**188**) which could both be acetylated in the 3-position in 91% and 96% yield, respectively, using a large excess of acetic anhydride (32 equiv.) (Scheme 5-12).[206]



Scheme 5-12: Regioselective acetylations as reported by Evtushenko.[206]

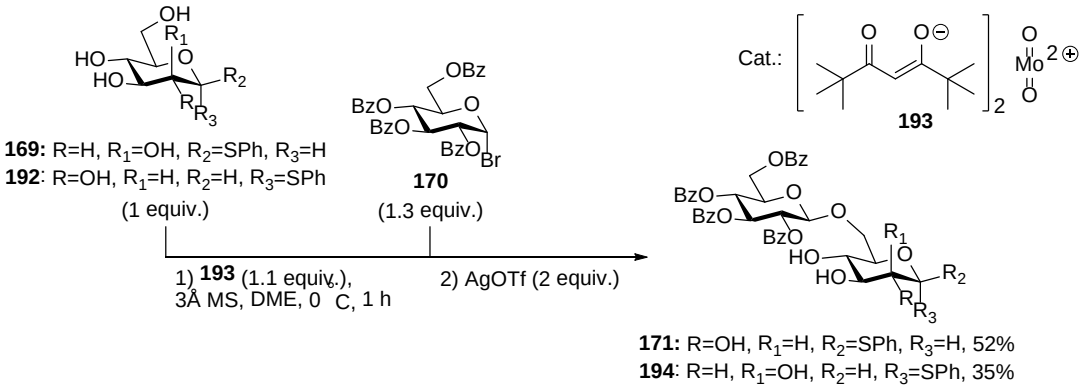
Later research of the same author demonstrated the ability of the molybdenum complex $MoO_2(acac)_2$ to mediate regioselective benzoylation of monosaccharides, of which the best result can be seen in Scheme 5-13.[207]



Scheme 5-13: Regioselective benzoylation using a molybdenum catalyst.[207]

Inspired by these results molybdenum was considered as possible catalytic candidate for regioselective glycosylations through coordination to the *O*-4 and *O*-6 positions. Gyrithe Lanz, a previous Ph.D. student in the group, performed the initial investigations.[221] Previous work in the group regarding regioselective glycosylations have been conducted utilising unprotected thiophenyl acceptors.[215,216,220] Hence, phenyl 1-thio-β-D-glucopyranoside (**169**) and phenyl 1-thio-α-D-mannopyranoside (**192**) were used as acceptors and perbenzoylated glucosyl

bromide (**170**) as the donor to explore the setup. The aim was to form the 1,6-linked product through activation of the primary alcohol by 4,6-complexation with Mo. Based on previous experience dimethoxyethane (DME) was used as the solvent and 3Å molecular sieves were added. Three different molybdenum complexes were investigated by G. Lanz; MoO_2Cl_2, $MoO_2(acac)_2$, and MoO_2(2,2,6,6-tetra-methyl-3,5-heptanedinate)$_2$ (**193**). The only molybdenum complex that seemed to promote the regioselective glycosylation was **193**. When the glucose derived donor **170** and acceptor **169** were used, the 1,6-linked product **171** was formed in 52% yield (Scheme 5-14). Exchanging the glucose acceptor **169** with the mannose acceptor **192** containing a 1,2-*cis* diol moiety lowered the yield of **194** to 35%. Hence, the presence of the 1,2-*cis* diol did not seem to enhance the selectivity.



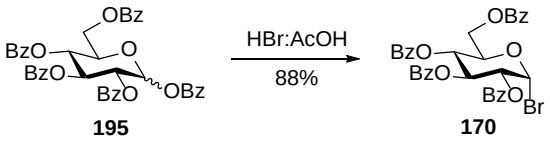
Scheme 5-14: Initial studies on regioselective glycosylations.[221]

5.3 Studies of molybdenum-mediated glycosylations

The results presented in this section represents research done by the author for this thesis.

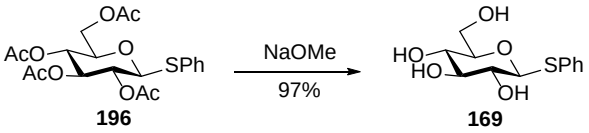
5.3.1 Optimisation with glucose-based acceptor

The per-benzoylated glucosyl bromide donor **170** was synthesised from perbenzoylated glucopyranose (**195**) according to literature procedures[221] (Scheme 5-15).



Scheme 5-15: Synthesis of donor **170**.

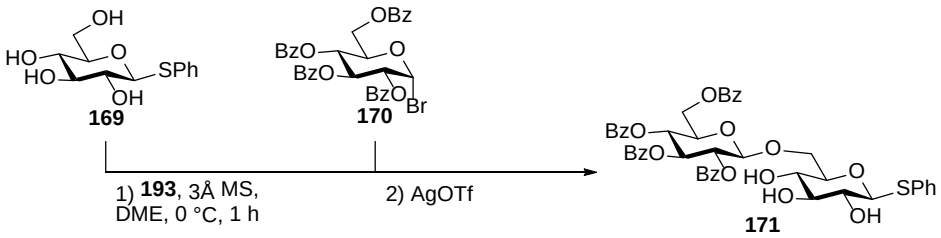
Acceptor **169** was obtained through deprotection of the acetyl groups in **196** following literature procedures[222] (Scheme 5-16).



Scheme 5-16: Zemplén deprotection of **169**.

Based on the preliminary studies performed with the molybdenum complex **193**, a series of studies were performed trying to optimise the preliminary investigations and thereby confirming the catalytic activity of the molybdenum complex in regioselective glycosylations (Table 5-1).

Table 5-1: Optimisation of the reaction conditions.



Entry:	Temperature (°C):	Reaction time (h):	Yield:[a]
1[b]	0 to r.t.	3	49%
2	0 to r.t.	3	47%[c]
3[d]	0 to r.t.	3	39%
4[e]	0 to r.t.	6	42%
5[f]	0 to r.t.	3	38%
6	-20	3	40%
7	-40	6	46%
8	-70	6	14%
9[g]	0 to r.t.	5.5	-
10[h]	0 to r.t.	3	40%
11[h]	-40	6	39%
12[h,i]	0 to r.t.	3	42%

Conditions: Acceptor **169** (1.0 equiv.), donor **170** (1.3 equiv.), MoO_2(2,2,6,6-tetra-methyl-3,5-heptanedinate)$_2$, (**193**) (1.1 equiv.), AgOTf (2.0 equiv.), DME, 3Å MS. [a]Isolated yield. [b]Old previously opened batch of the Mo catalyst. [c]Elimination found to be major by-product. [d]Added 3 equiv. of the donor. [e]Added 1.5 equiv. of AgOTf. [f]Concentration increased to the double. [g]Used Ag_2O (2.0 equiv.) as the promotor. [h]DME extra dried, 7.9 ppm. [i]No molybdenum catalyst added.

For the optimisation the same conditions as was defined by former Ph.D. student G. Lanz were utilised.[221] Firstly, it was found that the preliminary results could be recreated to a reasonable extend using both an old opened and a newly bought batch of the Mo complex **193**, obtaining **171** in 49% (Entry 1) and 47% yield (Entry 2), respectively. Furthermore, it was found that the major side-reaction was the elimination of the donor to the corresponding glycal (Entry 2). As

a consequence of the formation of the glycal, an increased amount of donor **170** (3.0 equiv.) was added to the reaction in Entry 3. It was hoped, that by adding more of the donor more of the desired product **171** could be obtained, but this was not the case. As the glycal most likely arise from the elimination of the α-triflate of the donor, less AgOTf was added in Entry 4, but approximately the same yield was obtained. Based on the previous experiments the formation of the glycal seemed to be considerably faster than the desired reaction, and it was therefore tried to increase the likelihood of the target reaction by doubling the concentration, but no improvement in the yield was observed (Entry 5). In the hope to change the reactivity, the temperature was tuned (Entry 6-8). Running the reaction at -20 °C lowered the yield minimally. On the other hand, running the reaction at -70 °C decreased the yield dramatically to 14%. With a reaction temperature of -40 °C, approximately the same yield was obtained as in Entry 2. Trying to use Ag_2O as the promotor as to avoid AgOTf and hence the α-triflate, meant that the reaction never started (Entry 9). At this point, it was realised that the DME used as solvent was not as dry as expected; and the solvent was therefore dried to a water content of 7.9 ppm. With the dried solvent, the reaction was run with reaction temperatures of 0 °C to r.t. (Entry 10) and at -40 °C (Entry 11), as it seemed that the previously best results were obtained at these temperatures. Having the solvent extra dry also did not seem to have any real influence on the yield, although, the dried solvent was utilised for the further optimisations.

During the optimisations, methyl α-D-glucopyranoside was also tried as the acceptor in the reaction at 0 °C to r.t. to clarify if the thiophenyl acceptor **169** was the limiting factor. However for this reaction the expected 1,6-linked disaccharide **197** was obtained in only 5% yield alongside the 1,2-linked product **198** (3%) and the 1,3-linked product **199** (2%) (Figure 5-1). The major product was the trisaccharide **200** with the per-benzoylated donor attached in both the 2- and 6-position in 18% yield. In addition, phenyl 1-thio-β-D-galactopyranoside was used as the acceptor, where the reaction was run at 0 °C to r.t. For this reaction the 1,6-linked product **201** was obtained in 45% yield and the 1,3-linked product **202** in 4% yield (Figure 5-1). This indicates a possible catalytic effect of molybdenum as the 1,3-product would normally be the favoured product.

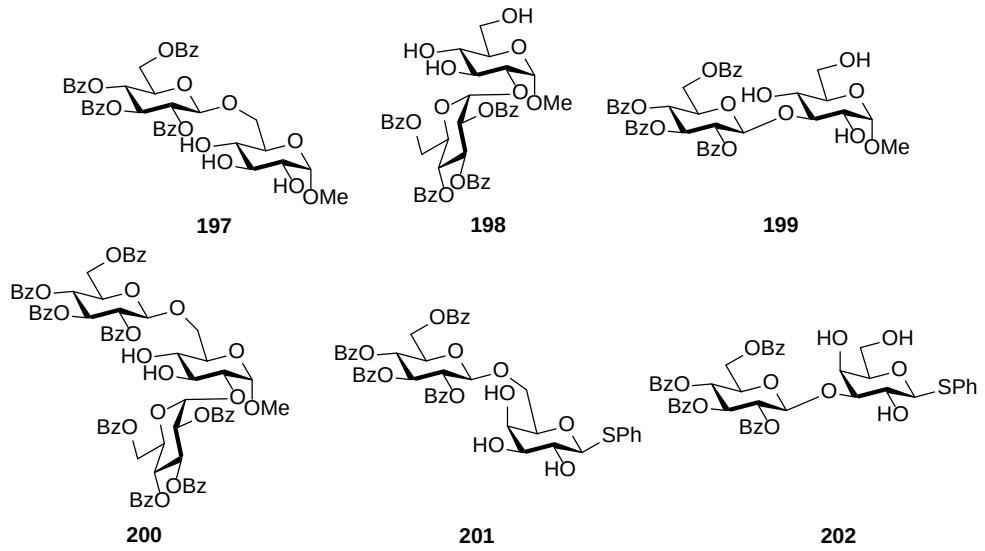
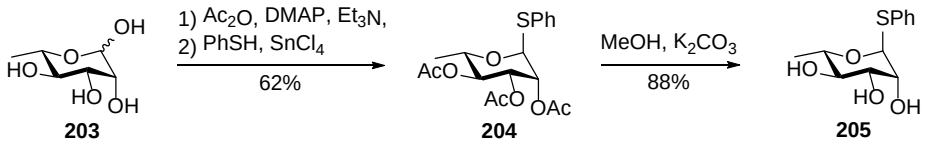


Figure 5-1: Products obtained using methyl α-D-glucopyranoside and phenyl 1-thio-β-D-galactopyranoside as acceptors using the same glycosylation protocol as previously described.

Seemingly, no changes in the reaction conditions could raise the yield of the glucose-based glycosylation to more than approximately 45%. Hence, as a control the reaction was run without the addition of the molybdenum complex **193** (Table 5-1 Entry 12). This reaction provided **171** in 42%, confirming that the molybdenum complex did not have any catalytic activity for the reaction, and the regioselectivity observed was caused by the inherent reactivity in the system.

5.3.2 Optimisation based on 1,2-*cis* diols

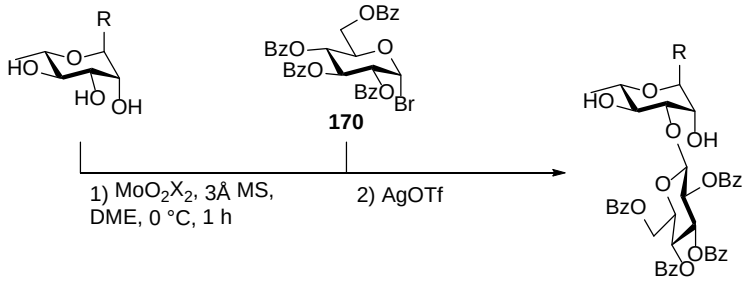
The previous section indicated that the molybdenum complex **193** does not have any catalytic activity in the former system targeting the 1,6-linked disaccharide where the acceptor is based on glucose. It was speculated that the lack of catalytic activity could be caused by the absence of 1,2-*cis* diols in the acceptor. This functionality was a needed for the regioselective protection group manipulation as described by E. Evtushenko.[206,207] Consequently, phenyl 1-thio-α-L-rhamnopyranoside (**205**) was synthesised from L-rhamnose (**203**) according to a literature procedure.[223] Firstly, acetyl protection groups was introduced, then the anomeric thiophenyl was installed in 62% yield. In the following reaction, the acetyl groups were again removed providing **205** in 88% yield (Scheme 5-18).



Scheme 5-17: Synthesis of phenyl 1-thio-α-L-rhamnopyranoside (**203**).

The glycosylation with the rhamnose acceptor **205** was run under the same conditions as the starting point for optimization based on the glucose acceptors (Table 5-1 entry 2). Due to the presence of the 1,2-*cis* diol in the acceptor, the expected product would be the 1,3-linked disaccharide, as coordination of the complex to the *O*-2 and *O*-3 should favor reaction on the most reactive hydroxyl, the equatorial. The reaction was run both at 0 °C to r.t. and at -40 °C resulting in the expected 1,3-linked product **206** in 33% in both cases but alongside the 1,4-linked product **209** in 20 and 15% yield, respectively (Table 5-2 entry 1 and 2 and Figure 5-2). Hence, the reactions indicated a slight preference for glycosylation on the 3-position of the acceptor, but it was not very pronounced, and furthermore, the overall yield was not great. As a last resort, the reaction was tried with the commercially available methyl α-L-rhamnopyranoside (**190**) as acceptor and trying the previously discarded molybdenum complex $MoO_2(acac)_2$ at 0 °C to r.t. and at -40 °C (Table 5-2, entry 3-4, Figure 5-2) as this was the Mo complex previously utilised for regioselective benzoylations. Under these conditions, the major product was the 1,3-linked product **207**. Due to difficulties with the purification, the yield of **207** is not precise, but the purification was still sufficient to determine it to be the major product. Furthermore for the reaction run at 0 °C to r.t. also 3% of the 1,2-linked product **208** and 8% of the 1,4-linked product **210** were isolated (Entry 3). For the reaction run at -40 °C the 1,2- and 1,4-linked products **208** and **210** were isolated in 6% and 8% yield respectively (Entry 4). The isolation of even the 1,2-linked product was surprising, and indicates that neither this molybdenum complex has catalytic activity in the glycosylation system.

Table 5-2: Reactions with rhamnose-based acceptors.



Entry:	R:	X:	Temperature (°C):	Time (h):	Product (yield):[a]
1	SPh	A	0 to r.t.	2.5	**206** (33%), **208** (20%)
2	SPh	A	-40	4	**206** (33%), **208** (15%)
3	OMe	B	0 to r.t.	Overnight	**207** (approx. 23%), **208** (3%), **210** (8%)
4	OMe	B	-40	2	**207** (30-40%), **208** (6%), **210** (8%)

Conditions: Acceptor (1.0 equiv.), donor **170** (1.3 equiv.), MoO_2X_2 (1.1 equiv., A = 2,2,6,6-tetra-methyl-3,5-heptanedinate, B = acac), AgOTf (2.0 equiv.), DME, 3Å MS. [a]Isolated yield.

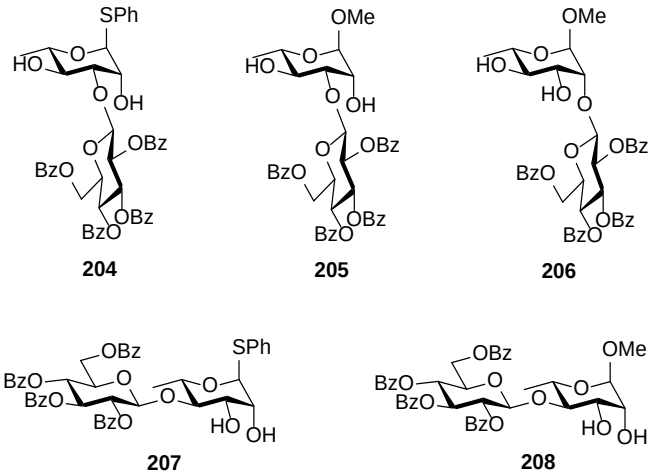


Figure 5-2: Products obtained during the investigations with rhamnose-based acceptors in Table 5-2.

5.4 Concluding remarks

For this project, regioselective glycosylations of unprotected carbohydrate acceptors promoted by a molybdenum complex was investigated. However, it was not possible to show any significant induction of regioselectivity in the systems utilised. For the glucose-based reaction it was even shown, that the yields obtained are most likely caused by the inherent regioselectivity in the system. Furthermore, it was investigated if the presence of 1,2-*cis* vicinal diols could cause coordination to the molybdenum complex, but the results did not show any

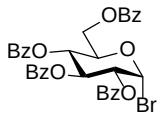
clear regioselectivity in this case neither. Hence, this project was ended with the knowledge that in this type of system molybdenum does not seem to have any catalytic activities.

5.5 Experimental

General considerations

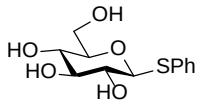
Starting materials, reagents, and solvents were purchased from commercial suppliers and were used without further purification unless otherwise noted. All solvents used were of analytical HPLC grade. Anhydrous solvents were obtained from an Innovative Technology PS-MD-7 PureSolv solvent purification system. Air- and moisture sensitive experiments were conducted under inert atmosphere (N_2) in dried glassware and anhydrous solvents. Water content was checked by 899 Coulometer from Metrohm. Crushed 3 Å molecular sieves for glycosylations were stored in the oven at 110 °C, before activation with heating under vacuum. Evaporation of solvents was performed with a VWR International Laborota 400 under reduced pressure (*in vacuo*) at temperatures between 36-50 °C. Thin-layer chromatography (TLC) was carried out using Merck Aluminium Sheets pre-coated with 0.25 mm silica gel, C-60 F_{254} plates. TLC plates were inspected under UV light or visualized by charring after dipping in a cerium ammoniumsulfate solution (1% cerium(IV)sulphate and 2.5% ammonium heptamolybdate in a 10% sulfuric acid solution). Flash column chromatography was performed using Merck Geduran silica gel 60 Å (40-63 µm) as the stationary phase. Eluent systems are specified for each R_f-value, and ratios are given as volume ratios. NMR spectre were recorded with a Bruker Ascend™ 400 (400 MHz and 101 MHz) with a Prodigy cyroprobe. The recorded chemical shift are reported in parts per million (δ = scale) relative to the residual solvent peak in $CDCl_3$ (δ_H = 7.26 ppm, δ_C = 77.16 ppm), and methanol-*d4* (δ_H = 4.87 ppm, δ_C = 49.00 ppm). The annotation (2C), (3C) ect. indicates when multiple carbon signals are overlapping. Coupling constants (*J*) are reported in Hz. Multiplicities are reported as singlet (s), broad singlet (bs), doublet (d), doublets of doublets (dd), doublets of triplets (dt), douplet of quartets (dq), doublets of doublets of doublets (ddd), triplet (t), and multiplet (m). Assignment of ^1H and ^{13}C resonances were based on APT, DQF-COSY, HSQC, H2BC, HMBC, and TOCSY when needed. Optical rotation was measured on a Perkin Elmer Model 341 Polarimeter. HRMS analysis was performed on either a UHPLC-QTOF system (Dionex ultimate 3000 and Bruker MaXis) with an electrospray ionization (ESI) source and controlled using Data Analysis 4.2 software. New compounds have been characterised by R_f, ^1H NMR, ^{13}C NMR or ^{13}C APT.

2,3,4,6-tetra-*O*-benzoyl-α-D-glucopyranosyl bromide (170)[221]



Pentabenzoate **195** (2.996 g, 4.28 mmol, 1.0 equiv.) was dissolved in CH_2Cl_2 (30.0 mL) and cooled to 0 °C. 33% HBr in acetic acid (5.0 mL, 28.14 mmol, 6.6 equiv.) was added and the reaction mixture was left to stir at r.t. after 10 min. After 3.5 h TLC (hexane/EtOAc 2:1) showed complete conversion to the product. Ice-cold water (30 mL) was added, and after 10 min of stirring the phases was separated. The organic phase was washed with $NaHCO_3$ and H_2O, dried over Na_2SO_4, filtered, and conc. *in vacuo*. The resultant residue was recrystallised from Et_2O obtaining the white crystalline product (2.490 g, 88%). $\mathbf{R_f}$(heptane/EtOAc 2:1) 0.50. $[\boldsymbol{\alpha}]_{\boldsymbol{D}}^{\boldsymbol{298\,K}}$ = +118.1 (*c* 1.0, $CHCl_3$). **^{1}H NMR** (400 MHz, $CDCl_3$) δ 8.09 – 8.06 (m, 2H, Ar*H*), 8.02 – 7.99 (m, 2H, Ar*H*), 7.97 – 7.94 (m, 2H, Ar*H*), 7.89 – 7.87 (m, 2H, Ar*H*), 7.60 – 7.50 (m, 3H, Ar*H*), 7.47 – 7.29 (m, 9H, Ar*H*), 6.87 (d, $J_{1,2}$ = 4.0 Hz, 1H, H1), 6.27 (t, J = 9.8 Hz, 1H, H3), 5.82 (t, J = 10.0 Hz, 1H, H4), 5.33 (dd, $J_{2,3}$ = 10.0 Hz, $J_{2,1}$ 4.0 Hz, 1H, H2), 4.74 (ddd, $J_{5,4}$ = 10.2 Hz, $J_{5,6'}$ = 4.3 Hz, $J_{5,6}$ = 2.7 Hz, 1H, H5), 4.67 (dd, $J_{6,6}$ = 12.5 Hz, $J_{6,5}$ = 2.7 Hz, 1H, H6), 4.52 (dd, $J_{6',6}$ = 12.5 Hz, $J_{6',5}$ = 4.5 Hz, 1H). **^{13}C NMR** (101 MHz, $CDCl_3$) δ 166.2 (1C, C_q), 165.7 (1C, C_q), 165.4 (1C, C_q), 165.2 (1C, C_q) (CO, Bz), 133.9 (1C, Ar*C*H), 133.8 (1C, Ar*C*H), 133.5 (1C, Ar*C*H), 133.4 (1C, Ar*C*H), 130.2 (1C, Ar*C*H), 130.1 (1C, Ar*C*H), 130.0 (1C, Ar*C*H), 129.9 (1C, Ar*C*H), 129.6 (C_q, Ar*C*), 128.9 (C_q, Ar*C*), 128.7 (1C, Ar*C*H), 128.7 (C_q, Ar*C*), 128.6 (1C, Ar*C*H), 128.6 (1C, Ar*C*H), 128.5 (C_q, Ar*C*), 128.5 (1C, Ar*C*H), 87.0 (C1), 72.9 (C5), 71.6 (C2), 70.8 (C3), 68.1 (C4), 62.1 (C6). NMR data are in accordance with literature values.[224]

Phenyl 1-thio-β-D-glucopyranoside (169)[222]

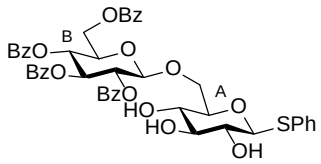


Tetraacetate **196** (4.998 g, 11.35 mmol, 1.0 equiv.) was dissolved in MeOH (50.0 mL) and NaOMe was added until the solution was basic. After 3.5 h TLC (heptane/EtOAc) showed complete conversion to the product. The reaction mixture was neutralized with Amberlite H^+, filtered, and conc. *in vacuo*. The residue was crystallised from EtOAc obtaining the product as a white powder (3.009 g, 97%). $\mathbf{R_f}$(EtOAc) 0.23. $[\boldsymbol{\alpha}]_{\boldsymbol{D}}^{\boldsymbol{298\,K}}$ = -56.2 (*c* 1.0, MeOH). **^{1}H NMR** (400 MHz, Methanol-d_4) δ 7.58 – 7.5 (m, 2H, ArH_{ortho}), 7.33 – 7.24 (m, 3H, ArH_{meta}, ArH_{para}), 4.61 (d, $J_{1,2}$ = 9.8 Hz, 1H, H1), 3.88 (dd, $J_{6,6'}$ = 12.1 Hz, $J_{6,5}$ = 1.8 Hz, 1H, H6), 3.68 (dd, $J_{6',6}$ = 12.1 Hz, $J_{6',5}$ = 5.3 Hz, 1H, H6'), 3.40 (t, $J_{3,2}$ = 8.7 Hz, 1H, H3), 3.36 – 3.28 (m, 2H, H4, H5), 3.23 (dd, $J_{2,1}$ = 9.7 Hz, $J_{2,3}$ = 8.7 Hz, 1H, H2). **^{13}C NMR** (101 MHz, Methanol-d_4) δ 135.3 (C_q, ArC_{ipso}), 132.7 (2C, ArC_{ortho}H), 129.8 (2C, ArC_{meta}H), 128.3 (ArC_{para}H),

89.4 (C1), 82.0 (C5), 79.6 (C3), 73.7 (C2), 71.3 (C4), 62.8 (C6). NMR data are in accordance with literature values.[225]

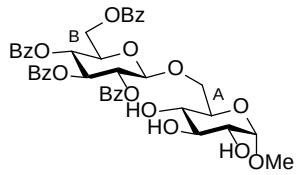
General procedure:[221] A Schlenk flask with crushed 3Å molecular sieves was placed under vacuum and heated. Afterwards an atm. of N_2 was applied. The acceptor (0.37 mmol, 1.0 equiv.) and the molybdenum complex (0.41 mmol, 1.1 equiv.) were added and dissolved in dry DME (3 mL). The mixture was cooled to 0 °C and stirred for 1 h. At this point, the donor (0.48 mmol, 1.3 equiv.) and the promotor (0.55 mmol, 2.0 equiv.) were added. The mixture was stirred at the temperature indicated until TLC showed complete consumption of the acceptor (1-24 h). The mixture was filtered through a pad of Celite and conc. *in vacuo*. The crude product was purified by flash column chromatography (0-3% MeOH in CH_2Cl_2).

Phenyl 2,3,4,6-tetra-*O*-benzoyl-β-D-glucopyranosyl-(1→6)-1-thio-β-D-glucopyranoside (171)



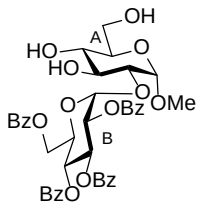
R_f(CH_2Cl_2, 4% MeOH) 0.47. $[\alpha]_D^{298\,K}$ = +3.3 (*c* 1.0, $CHCl_3$). **^{1}H NMR** (400 MHz, $CDCl_3$) δ 8.04 – 8.02 (m, 2H, Ar*H*), 7.92 – 7.89 (m, 4H, Ar*H*), 7.84 – 7.82 (m, 2H, Ar*H*), 7.55 – 7.28 (m, 17H, Ar*H*), 5.89 (t, *J* = 9.6 Hz, 1H, $H3^B$), 5.69 (t, *J* = 9.7 Hz, 1H, $H4^B$), 5.52 (dd, $J_{2B,3B}$ = 9.7 Hz, $J_{2B,1B}$ = 7.9 Hz, 1H, $H2^B$), 4.96 (d, $J_{1B,2B}$ = 7.9 Hz, 1H, $H1^B$), 4.71 (dd, $J_{6B,6'B}$ = 12.2 Hz, $J_{6B,5B}$ = 3.1 Hz, 1H, $H6^B$), 4.47 – 4.43 (m, 2H, $H1^A$, $H6'^B$), 4.14 – 4.08 (m, 2H, $H6^A$, $H5^B$), 3.91 (dd, $J_{6A,6'A}$ = 11.4 Hz, $J_{6A,5A}$ = 4.7 Hz, 1H, $H6'^A$), 3.50 – 3.39 (m, 3H, $H3^A$, $H4^A$, $H5^A$), 3.29 – 3.24 (m, 1H, $H2^A$), 2.64 (bs, 4H, OH). **^{13}C NMR** (101 MHz, $CDCl_3$) δ 166.5 (1C, C_q), 165.9 (1C, C_q), 165.4 (1C, C_q), 165.3 (1C, C_q) (CO, Bz), 133.6 (1C, Ar*C*H), 133.5 (1C, Ar*C*H), 133.4 (2C, Ar*C*H), 132.7 (2C, Ar*C*H), 132.1 (C_q, Ar*C*), 130.0 (4C, Ar*C*H), 129.9 (2C, Ar*C*H), 129.6 (C_q, Ar*C*), 129.2 (2C, Ar*C*H), 129.2 (C_q, Ar*C*), 128.9 (C_q, Ar*C*), 128.9 (C_q, Ar*C*), 128.6 (2C, Ar*C*H), 128.6 (2C, Ar*C*H), 128.5 (2C, Ar*C*H), 128.4 (2C, Ar*C*H), 128.3 (1C, Ar*C*H), 101.5 ($C1^B$), 88.0 ($C1^A$), 78.8 ($C5^A$), 77.8 ($C3^A$), 72.9 ($C3^B$), 72.5 ($C5^B$), 72.0 ($C2^A$/$C2^B$), 71.9 ($C2^A$/$C2^B$), 70.6 ($C4^A$), 69.6 ($C4^B$), 69.5 ($C6^A$), 62.9 ($C6^B$). **HRMS:** calcd. for $C_{46}H_{42}O_{14}SNa^+$ *m/z* 873.2187; found *m/z* 873.2178. NMR data are in accordance with literature values.[215]

Methyl 2,3,4,6-tetra-*O*-benzoyl-β-D-glucopyranosyl-(1→6)-α-D-glucopyranoside (197)



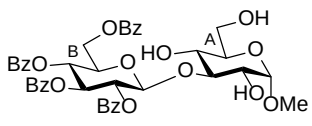
^{1}H NMR (400 MHz, $CHCl_3$) δ 8.04 – 8.02 (m, 2H, Ar*H*), 7.96 – 7.94 (m, 2H, Ar*H*), 7.91 – 7.89 (m, 2H, Ar*H*), 7.84 – 7.82 (m, 2H, Ar*H*), 7.57 – 7.28 (m, 12H, Ar*H*), 5.92 (t, *J* = 9.6 Hz, 1H, H3^B), 5.69 (t, *J* = 9.7 Hz, 1H, H4^B), 5.55 (dd, $J_{2B,3B}$ = 9.7 Hz, $J_{2B,1B}$ = 7.9 Hz, 1H, H2^B), 4.97 (d, $J_{1B,2B}$ = 7.9 Hz, 1H, H1^B), 4.67 (dd, $J_{6B,6B'}$ = 12.2 Hz, $J_{6B,5B}$ = 3.1 Hz, 1H, H6^B), 4.58 (d, *J* = 3.8 Hz, 1H, H1^A), 4.51 (dd, $J_{6B',6B}$ = 12.2 Hz, $J_{6B',5B}$ = 5.0 Hz, 1H, H6^B'), 4.20 – 4.14 (m, 2H, H6^A, H5^B), 3.83 (dd, *J* = 11.1, 5.5 Hz, 1H, H6^A'), 3.68 – 3.64 (m, 1H, H5^A), 3.58 (t, *J* = 9.2 Hz, 1H, H3^A), 3.42 – 3.28 (m, 2H, H2^A, H4^A), 3.20 (s, 3H, OC*H*$_3$), 2.25 (bs, 3H, OH). **APT NMR** (101 MHz, $CDCl_3$) δ 166.4 (1C, C_q), 165.9 (1C, C_q), 165.3 (2C, C_q) (CO, Bz), 133.6 - 128.4 (Ar*C*H, Ar*C*), 101.8 (C1^B), 99.1 (C1^A), 74.8 (C3^A), 72.9 (C3^B), 72.6 (C5^B), 72.3 (C4^A), 71.9 (C2^B), 70.5, 70.4 (C2^A/C5^A), 69.7 (C4^B), 69.2 (C6^A), 63.1 (C6^B), 55.3 (O*C*H$_3$).

Methyl 2,3,4,6-tetra-*O*-benzoyl-β-D-glucopyranosyl-(1→2)-α-D-glucopyranoside (198)



^{1}H NMR (400 MHz, $CHCl_3$) δ 8.09 – 8.03 (m, 2H, Ar*H*), 7.98 – 7.89 (m, 4H, Ar*H*), 7.84 – 7.79 (m, 2H, Ar*H*), 7.61 – 7.27 (m, 12H, Ar*H*), 5.94 (t, *J* = 9.7 Hz, 1H, H3^B), 5.65 (t, *J* = 9.7 Hz, 1H, H4^B), 5.56 (dd, $J_{2B,3B}$ = 9.8, $J_{2B,1B}$ = 7.9 Hz, 1H, H2^B), 5.24 (d, $J_{1B,2B}$ = 7.9 Hz, 1H, H1^B), 4.77 – 4.73 (m, 2H, H1^A, H6^B), 4.43 (dd, *J* = 12.2, 5.9 Hz, 1H, H6^B'), 4.20 (ddd, *J* = 9.7, 6.0, 2.9 Hz, 1H, H5^B), 3.92 – 3.88 (m, 1H, H3^A), 3.81 (dd, *J* = 11.8, 3.5 Hz, 1H, H6^A), 3.75 (dd, *J* = 11.8, 4.5 Hz, 1H, H6^A'), 3.63 – 3.59 (m, 1H, H5^A), 3.56 – 3.47 (m, 2H, H2^A, H4^A), 3.28 (s, 3H, OC*H*$_3$), 1.88 (bs, 3H, OH). **APT NMR** (101 MHz, $CHCl_3$) δ 166.2 (1C, C_q), 165.9 (1C, C_q), 165.4 (1C, C_q), 165.3 (1C, C_q) (CO, Bz), 133.7 - 128.5 (Ar*C*H, Ar*C*), 101.2 (C1^B), 99.5 (C1^A), 81.0 (C2^A), 72.8 (C3^B/C5^B), 72.7 (C3^B/C5^B), 72.3 (C2^B), 71.7 (C3^A), 70.8 (C4^A), 70.6 (C5^A), 69.5 (C4^B), 63.0 (C6^B), 62.5 (C6^A), 55.4 (O*C*H$_3$). NMR data are in accordance with literature values.[226]

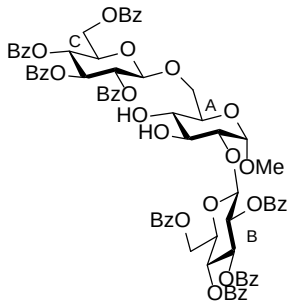
Methyl 2,3,4,6-tetra-*O*-benzoyl-β-D-glucopyranosyl-(1→3)-α-D-glucopyranoside (199)



^{1}H NMR (400 MHz, $CHCl_3$) δ 8.10 – 8.08 (m, 2H, Ar*H*), 7.97 – 7.92 (m, 4H, Ar*H*), 7.84 – 7.81 (m, 2H, Ar*H*), 7.60 – 7.27 (m, 12H, Ar*H*), 5.93 (t, *J* = 9.7 Hz, 1H, H3^B), 5.62 (t, *J* = 9.7 Hz, 1H, H4^B), 5.55 (dd, $J_{2B,3B}$ = 9.8 Hz, $J_{2B,1B}$ = 8.0 Hz, 1H, H2^B), 5.02 (d, $J_{1B,2B}$ = 8.0 Hz, 1H, H1^B), 4.78 (dd, $J_{6B,6B'}$ = 12.3, $J_{6B,5B}$ = 2.5 Hz, 1H, H6^B), 4.68 (d, *J* = 3.8 Hz, 1H, H1^A), 4.38 (dd, $J_{6B',6B}$ = 12.3 Hz, $J_{6B',5B}$ = 6.8 Hz, 1H, H6^B'), 4.24

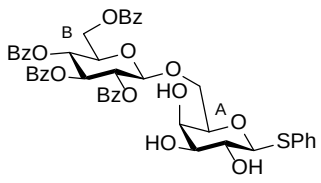
(ddd, J = 9.6, 6.8, 2.5 Hz, 1H, H5^{B}), 3.84 (dd, J = 11.6, 3.8 Hz, 1H, H6^{A}), 3.77 – 3.69 (m, 2H, H3^{A}, H6^{A}'), 3.63 – 3.58 (m, 1H, H5^{A}), 3.56 – 3.50 (m, 2H, H2^{A}, H4^{A}), 3.38 (s, 3H, OCH_3). **APT** NMR (101 MHz, $CHCl_3$) δ 166.3 (1C, C_q), 165.8 (1C, C_q), 165.6 (1C, C_q), 165.4 (1C, C_q) (CO, Bz), 133.8 - 128.5 (ArCH, ArC), 102.7 (C1^{B}), 99.1 (C1^{A}), 87.4 (C3^{A}), 73.0 (C5^{B}), 72.6 (C3^{B}), 72.2 (C2^{B}), 71.1 (C2^{A}/C5^{A}), 70.9 (C2^{A}/C5^{A}), 69.8 (C4^{A}), 69.4 (C4^{B}), 63.1 (C6^{A}/C6^{B}), 63.0 (C6^{A}/C6^{B}), 55.3 (OCH$_3$). NMR data are in accordance with literature values.[226]

Methyl 2,3,4,6-tetra-*O*-benzoyl-β-D-glucopyranosyl-(1→2)-[(1→6)-2,3,4,6-tetra-*O*-benzoyl-β-D-glucopyranosyl]-α-D-glucopyranoside (200)



^{1}H NMR (400 MHz, $CHCl_3$) δ 8.04 – 8.01 (m, 4H, ArH), 7.96 – 7.89 (m, 8H, ArH), 7.84 – 7.82 (m, 4H, ArH), 7.58 – 7.27 (m, 24H, ArH), 5.93 – 5.87 (m, 2H, H3^{B}, H3^{C}), 5.68 (t, J = 9.7 Hz, 1H, H4^{C}), 5.63 (t, J = 9.7 Hz, 1H, H4^{B}), 5.56 – 5.49 (m, 2H, H2^{B}, H2^{C}), 5.11 (d, J = 7.9 Hz, 1H, H1^{B}), 4.92 (d, J = 7.8 Hz, 1H, H1^{C}), 4.71 (dd, J = 12.2, 2.8 Hz, 1H, H6^{B}), 4.66 (dd, J = 12.2, 3.3 Hz, 1H, H6^{C}), 4.60 (d, J = 3.4 Hz, 1H, H1^{A}), 4.50 (dd, J = 12.2, 5.0 Hz, 1H, H6^{C}'), 4.39 (dd, J = 12.3, 5.8 Hz, 1H, H6^{B}'), 4.17 – 4.10 (m, 3H, H6^{A}, H5^{B}, H5^{C}), 3.81 – 3.76 (m, 2H, H3^{A}, H6^{A}'), 3.65 (ddd, J = 10.1, 5.2, 1.9 Hz, 2H, H5^{A}), 3.39 – 3.34 (m, 2H, H2^{A}, H4^{A}), 3.08 (s, 3H, OCH_3). **APT NMR** (101 MHz, $CHCl_3$) δ 166.3 (C_q), 166.2 (C_q), 165.9 (C_q), 165.9 (C_q), 165.3 (C_q), 165.3 (2C, C_q), 165.2 (C_q) (CO, Bz), 133.7 - 128.4 (ArCH, ArC), 101.8 (C1^{C}), 101.1 (C1^{B}), 99.2 (C1^{A}), 80.9 (C2^{A}), 72.9 (C3^{B}/C3^{C}), 72.8 (C3^{B}/C3^{C}), 72.6 (C5^{B}/C5^{C}), 72.4 (C5^{B}/C5^{C}), 72.2 (C2^{B}/C2^{C}), 72.0 (C2^{B}/C2^{C}), 71.5 (C3^{A}), 70.3 (C4^{A}), 70.0 (C5^{A}/C4^{C}), 69.8 (C5^{A}/C4^{C}), 69.4 (C4^{B}), 69.0 (C6^{B}/C6^{A}), 63.2 (C6^{B}/C6^{C}), 63.0 (C6^{B}/C6^{C}), 55.2 (OCH$_3$).

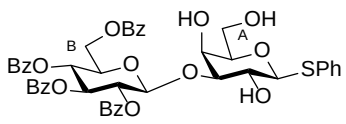
Phenyl 2,3,4,6-tetra-*O*-benzoyl-β-D-glucopyranosyl-(1→6)-1-thio-β-D-galactopyranoside (201)



R$_f$(CH_2Cl_2, 3% MeOH) 0.32. $[\alpha]_D^{298\,K}$ = +4.7 (c 1.0, $CHCl_3$). **^{1}H NMR** (400 MHz, $CDCl_3$) δ 8.09 – 8.07 (m, 2H, ArH), 7.93 – 7.90 (m, 4H, ArH), 7.83 – 7.81 (m, 2H, ArH), 7.60 – 7.28 (m, 17H, ArH), 5.88 (t, J = 9.7 Hz, 1H, H3^{B}), 5.68 (t, J = 9.7 Hz, 1H, H4^{B}), 5.48 (dd, $J_{H2B,H3B}$ = 9.8 Hz, $J_{H2B,H1B}$ = 7.9 Hz, 1H, H2^{B}), 4.94 (d, $J_{H1B,H2B}$ = 7.9 Hz, 1H, H1^{B}), 4.82 (dd, $J_{H6B,H6B'}$ = 12.3 Hz, $J_{H6B,H5B}$ = 2.9 Hz, 1H, H6^{B}), 4.44 – 4.38 (m, 2H, H1^{A}, H6^{B}'), 4.14 (ddd, $J_{H5B,H4B}$ = 9.8 Hz, $J_{H5B,H6B'}$ = 4.7 Hz, $J_{H5B,H6B}$ =

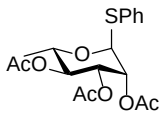
3.0 Hz, 1H, H5^{B}), 4.06 (dd, $J_{H6A,H6A'}$ = 10.3 Hz, $J_{H6A,H5A}$ = 7.1 Hz, 1H, H6^{A}), 3.98 – 3.94 (m, 2H, H4^{A}, H6$^{A'}$), 3.65 – 3.60 (m, 2H, H2^{A}, H5^{A}), 3.49 (dd, $J_{H3A,H2A}$ = 9.0 Hz, $J_{H3A,H4A}$ = 3.5 Hz, 1H, H3^{A}). **^{13}C NMR** (101 MHz, $CDCl_3$) δ 166.5 (C_q), 165.9 (C_q), 165.3 (C_q), 165.3 (C_q) (CO, Bz), 133.7 (1C, Ar*C*H), 133.6 (1C, Ar*C*H), 133.5 (1C, Ar*C*H), 133.4 (1C, Ar*C*H), 132.5 (C_q Ar*C*), 132.5 (2C, Ar*C*H), 130.0 (2C, Ar*C*H), 130.0 (2C, Ar*C*H), 130.0 (2C, Ar*C*H), 129.9 (2C, Ar*C*H), 129.4 (C_q Ar*C*), 129.2 (2C, Ar*C*H), 128.9 (C_q Ar*C*), 128.8 (C_q Ar*C*), 128.7 (2C, Ar*C*H), 128.6 (2C, Ar*C*H), 128.6 (2C, Ar*C*H), 128.5 (2C, Ar*C*H), 128.1 (1C, Ar*C*H), 101.3 (C1^{B}), 88.7 (C1^{A}), 77.4 (C5^{A}), 74.4 (C3^{A}), 72.9 (C3^{B}/C5^{B}), 72.7 (C3^{B}/C5^{B}), 71.8 (C2^{B}), 70.2 (C2^{A}), 69.4 (C4^{B}), 68.0 (C4^{A}), 67.9 (C6^{A}), 62.6 (C6^{B}). **HRMS:** calcd. for $C_{46}H_{42}O_{14}SNa^+$ *m/z* 873.2187; found *m/z* 873.2181. Data are in accordance with literature.[215]

Phenyl 2,3,4,6-tetra-*O*-benzoyl-β-D-glucopyranosyl-(1→3)-1-thio-β-D-galactopyranoside (202)



^{1}H NMR (400 MHz, $CDCl_3$) δ 8.10 – 7.82 (m, 8H, Ar*H*), 7.59 – 7.23 (m, 12H, Ar*H*), 5.94 (t, *J* = 9.7 Hz, 1H, H3^{B}), 5.63 (t, *J* = 9.7 Hz, 1H, H4^{B}), 5.49 (dd, $J_{2B,3B}$ = 9.8 Hz, $J_{2B,1B}$ = 7.9 Hz, 1H, H2^{B}), 5.15 (d, $J_{1B,2B}$ = 7.9 Hz, 1H, H1^{B}), 4.68 (dd, $J_{6B,6'B}$ = 12.2 Hz, $J_{6B,5B}$ = 2.9 Hz, 1H, H6^{B}), 4.49 (dd, $J_{6'B,6B}$ = 12.2 Hz, $J_{6'B,5B}$ = 6.1 Hz, 1H, H6$^{B'}$), 4.45 (d, $J_{1A,2A}$ = 9.7 Hz, 1H, H1^{A}), 4.18 (ddd, $J_{5B,4B}$ = 9.2 Hz, $J_{5B,6'B}$ = 6.0 Hz, $J_{5B,6B}$ = 2.9 Hz, 1H, H5^{B}), 4.05 (d, *J* = 3.1 Hz, 1H, H4^{A}), 3.85 (dd, $J_{6A,6'A}$ = 11.8 Hz, $J_{6A,5A}$ = 6.8 Hz, 1H, H6^{A}), 3.74 (t, *J* = 9.3 Hz, 1H, H2^{A}), 3.66 – 3.55 (m, 2H, H3^{A}, H6$^{A'}$), 3.45 (m, 1H, H5^{A}). **^{13}C NMR** (101 MHz, $CDCl_3$) δ 166.3 (1C, C_q), 165.9 (1C, C_q), 165.7 (1C, C_q), 165.4 (1C, C_q), 133.8 - 128.0 (Ar*C*H, Ar*C*), 101.7 (C1^{B}), 88.3 (C1^{A}), 83.8 (C3^{A}), 78.2 (C5^{A}), 72.7 (C5^{B}), 72.5 (C3^{B}), 72.4 (C2^{B}), 69.6 (C4^{B}), 68.8 (C4^{A}), 68.5 (C2^{A}), 62.9 (C6^{B}), 62.6 (C6^{A}). Spectre contain traces of **199**.

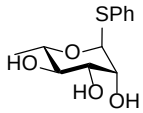
Phenyl 2,3,4-tri-*O*-acetyl-1-thio-α-L-rhamnopyranoside (204)[223]



L-Rhamnose monohydrate (**203**) (2.992 g, 18.23 mmol, 1.0 equiv.) was suspended in CH_2Cl_2 (15 mL) and cooled to 0 °C. Et_3N (20.0 mL, 143.49 mmol, 7.8 equiv.), DMAP (0.223 g, 1.83 mmol, 0.1 equiv.) and Ac_2O (10.0 mL, 105.79 mmol, 5.8 equiv.) were added at 0 °C. After 10 min the reaction mixture was stirred at r.t. for 6 h at which point TLC (EtOAc) showed full consumption of the starting material, and another TLC (hexane/EtOAc 2:1) showed formation of the product. MeOH (0.5 mL) was added and the

mixture was stirred for a few minutes after which it was conc. *in vacuo*. EtOAc was added and the mixture was washed with brine. The aqueous phase was extracted with EtOAc. The combined organic phases were washed with 1 M HCl, $NaHCO_3$ (until pH = 7), brine, dried over $MgSO_4$, filtered, and conc. *in vacuo*. The crude (6.050 g) was dissolved in CH_2Cl_2 (85 mL) and cooled to 0 °C. PhSH (2.0 mL, 19.48 mmol, 1.1 equiv.) and $SnCl_4$ 1 M in heptane (12.5 mL, 12.5 mmol, 0.7 equiv.) were added. After 6.5 h TLC (hexane/EtOAc 2:1) showed completion of the reaction with formation of two products. Et_2O (350 mL) was added and the mixture was washed with 1 M HCl, H_2O, $NaHCO_3$, and brine, dried over $MgSO_4$, filtered, and conc. *in vacuo*. The residue was purified by flash column chromatography (heptane/EtOAc 7:2). The α-product **204** was obtained as a white solid (4.345 g, 62%) and in addition an α/β mixture (1.513 g, 22%) (not characterised). $\mathbf{R_f}$(hexane/EtOAc 2:1) 0.59. **^{1}H NMR** (400 MHz, $CDCl_3$) δ 7.48 – 7.45 (m, 2H, ArH_{ortho}), 7.33 – 7.25 (m, 3H, ArH_{meta}, ArH_{para}) 5.49 (dd, $J_{H2,H3}$ = 3.3 Hz, $J_{H2,H1}$ = 1.6 Hz, 1H, H2), 5.40 (d, $J_{H1,H2}$ = 1.4 Hz, 1H, H1), 5.28 (dd, $J_{H3,H4}$ = 10.1 Hz, $J_{H3,H2}$ = 3.3 Hz, 1H, h3), 5.14 (t, J = 9.9 Hz, 1H, H4), 4.36 (dq, $J_{H5,H4}$ = 9.7 Hz, $J_{H5,H6}$ = 6.2 Hz, 1H, H5), 2.14 (s, 3H, COCH_3), 2.07 (s, 3H, COCH_3), 2.01 (s, 3H, COCH_3), 1.24 (d, $J_{H6,H5}$ = 6.2 Hz, 1H, H6). **^{13}C NMR** (101 MHz, $CDCl_3$) δ 170.1 (C_q), 170.1 (C_q), 170.0 (C_q) (CO, Ac), 133.4 (C_q, C_{ipso}), 131.9 (2C, C_{ortho}), 129.3 (2C, C_{meta}), 128.0 (1C, C_{para}), 85.8 (C1), 71.4 (C2), 71.3 (C4), 69.5 (C3), 67.9 (C5), 21.0 (CO*C*H_3), 20.9 (CO*C*H_3), 20.8 (CO*C*H_3), 17.5 (C6). NMR data are in accordance with literature values.[227,228]

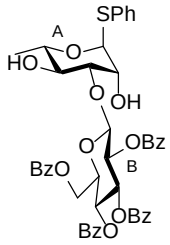
Phenyl 1-thio-α-L-rhamnopyranoside (205)[223]



Triacetate **204** (4.274 g, 11.18 mmol, 1.0 equiv.) dissolved in MeOH/THF 1:1 (54.0 mL) was mixed with K_2CO_3 (0.309 g, 2.23 mmol, 0.2 equiv.) and left to stir at r.t. After 5 h TLC (hexane/EtOAc 2:1) showed consumption of the starting material and another TLC (CH_2Cl_2, 10% MeOH) showed formation of a product. The mixture was filtered through a pad of Celite, conc. *in vacuo*, and purified by flash column chromatography (4-6% MeOH in CH_2Cl_2). The product was a white crystalline solid (2.520 g, 88%). $\mathbf{R_f}$(CH_2Cl_2, 10% MeOH) 0.65. $[\boldsymbol{\alpha}]_{\boldsymbol{D}}^{\boldsymbol{298\,K}}$ = -262.6 (*c* 1.0, $CHCl_3$). **^{1}H NMR** (400 MHz, CD_3OD) δ 7.49 – 7.46 (m, 2H, ArH_{ortho}), 7.35 – 7.25 (m, 3H, ArH_{meta}, ArH_{para}), 5.40 (d, $J_{1,2}$ = 1.3 Hz, 1H, H1), 4.10 (dd, $J_{2,3}$ = 3.3 Hz, $J_{2,1}$ = 1.5 Hz, 1H, H2), 4.09 – 4.03 (m, 1H, H5), 3.67 (dd, $J_{3,4}$ = 9.5 Hz, $J_{3,2}$ = 3.3 Hz, 1H, H3), 3.48 (t, J = 9.4 Hz, 1H, H4), 1.28 (d, $J_{6,5}$ = 6.2 Hz, 3H, H6). **^{13}C NMR** (101 MHz, CD_3OD) δ 136.0 (C_q, ArC_{ipso}), 132.6 (2C, ArC_{meta}H), 130.0 (2C, ArC_{orthof}H), 128.4 (1C, ArC_{para}H), 90.1 (C1), 74.1

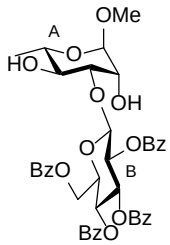
(C4), 73.8 (C2), 72.9 (C3), 70.9 (C5), 17.8 (C6). **HRMS:** calcd. for $C_{12}H_{16}O_4SNa^+$ *m/z* 279.0662; found *m/z* 279.0657.

Phenyl 2,3,4,6-tetra-*O*-benzoyl-β-D-glucopyranosyl-(1→3)-1-thio-α-L-rhamnopyranoside (206)



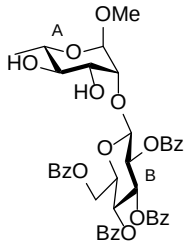
^{1}H NMR (400 MHz, $CDCl_3$) δ 8.10 – 8.06 (m, 2H, Ar*H*), 7.97 – 7.94 (m, 4H, Ar*H*), 7.85 – 7.83 (m, 2H, Ar*H*), 7.54 – 7.24 (m, 17H, Ar*H*), 6.00 (t, *J* = 9.7 Hz, 1H, $H3^B$), 5.63 (t, *J* = 9.8 Hz, 1H, $H4^B$), 5.55 (dd, $J_{2B,3B}$ = 9.9 Hz, $J_{2B,1B}$ = 7.9 Hz, 1H, $H2^B$), 5.35 – 5.30 (m, 1H, $H1^A$), 5.06 (d, $J_{1B,2B}$ = 7.9 Hz, 1H, $H1^B$), 4.72 (dd, $J_{6B,6B'}$ = 12.2 Hz, $J_{6B,5B}$ = 2.6 Hz, 1H, $H6^B$), 4.51 (dd, $J_{6B',6B}$ = 12.2 Hz, $J_{6B',5B}$ = 6.9 Hz, 1H, $H6^{B'}$), 4.30 – 4.25 (m, 2H, $H2^A$, $H5^B$), 4.09 (dd, $J_{5A,4A}$ = 9.2 Hz, $J_{5A,6A}$ = 6.2 Hz, 1H, $H5^A$), 3.77 (dd, $J_{3A,4A}$ = 9.1 Hz, $J_{3A,2A}$ = 3.0 Hz, 1H, $H3^A$), 3.69 (t, *J* = 9.2 Hz, 1H, $H4^A$), 1.24 (d, $J_{6A,5A}$ = 6.2 Hz, 3H, $H6^A$). **APT NMR** ^{13}C NMR (101 MHz, $CDCl_3$) δ 166.4 (1C, C_q), 165.8 (2C, C_q), 165.4 (1C, C_q) (CO, Bz), 134.3 - 127.3 (Ar*C*H, Ar*C*), 101.9 ($C1^B$), 87.0 ($C1^A$), 83.6 ($C2^A$), 72.9 ($C5^B$), 72.5, 72.4 ($C2^B$, $C3^B$), 71.3, 71.2 ($C2^A$, $C4^A$), 69.5 ($C4^B$), 68.8 ($C5^A$), 62.8 ($C6^B$), 17.5 ($C6^A$). Spectra contain small impurities and solvent peaks.

Methyl 2,3,4,6-tetra-*O*-benzoyl-β-D-glucopyranosyl-(1→3)-α-L-rhamnopyranoside (207)



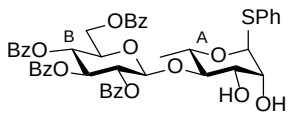
^{1}H NMR (400 MHz, $CDCl_3$) δ 8.09 – 8.06 (m, 2H, Ar*H*), 7.95 – 7.92 (m, 4H, Ar*H*), 7.85 – 7.82 (m, 2H, Ar*H*), 7.60 – 7.27 (m, 12H, Ar*H*), 5.95 (t, *J* = 9.7 Hz, 1H, $H3^B$), 5.61 (t, *J* = 9.8 Hz, 1H, $H4^B$), 5.51 (dd, $J_{2B,3B}$ = 9.8 Hz, $J_{2B,1B}$ = 7.9 Hz, 1H, $H2^B$), 4.99 (d, $J_{1B,2B}$ = 7.9 Hz, 1H, $H1^B$), 4.72 (dd, $J_{6B,6B'}$ = 12.2 Hz, $J_{6B,6B'}$ = 2.8 Hz, 1H, $H6^B$), 4.49 – 4.45 (m, 2H, $H1^A$, $H6^{B'}$), 4.19 (ddd, $J_{5B,4B}$ = 9.6 Hz, $J_{5B,6B'}$ = 6.8 Hz, $J_{5B,H6B}$ = 2.9 Hz, 1H, $H5^B$), 3.98 (dd, $J_{2A,3A}$ = 3.2 Hz, $J_{2A,1A}$ = 1.7 Hz, 1H, $H2^A$), 3.72 (dd, $J_{3A,4A}$ = 8.7 Hz, $J_{3A,2A}$ = 3.2 Hz, 1H, $H3^A$), 3.62 – 3.54 (m, 2H, $H4^A$, $H5^A$), 3.24 (s, 3H, OCH_3), 1.25 (d, $J_{6A,5A}$ = 5.7 Hz, 3H, $H6^A$). **APT NMR** (101 MHz, $CDCl_3$) δ 165.8 (1C, C_q), 165.8 (1C, C_q), 165.4 (1C, C_q) (CO, Bz), 133.8 - 128.5 (Ar*C*H, Ar*C*), 101.9 ($C1^B$), 100.4 ($C1^A$), 83.6 ($C3^A$), 72.8 ($C5^B$), 72.5, 72.5 ($C2^B$, $C3^B$), 70.9 ($C4^A$), 69.6, 69.5 ($C2^A$, $C4^B$), 67.4 ($C5^A$), 62.8 ($C6^B$), 54.9 (O*C*H_3), 17.7 ($C6^A$). The spectra contain small impurities and solvent peaks.

Methyl 2,3,4,6-tetra-*O*-benzoyl-β-D-glucopyranosyl-(1→2)-α-L-rhamnopyranoside (208)



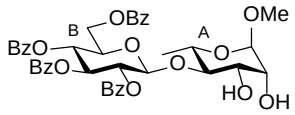
^{1}H NMR (400 MHz, $CDCl_3$) δ 8.03 – 8.01 (m, 2H, Ar*H*), 7.98 – 7.96 (m, 2H, Ar*H*), 7.92 – 7.89 (m, 2H, Ar*H*), 7.84 – 7.82 (m, 2H, Ar*H*), 7.59 – 7.27 (m, 12H, Ar*H*), 5.92 (t, J = 9.6 Hz, 1H, H3^B), 5.65 – 5.56 (m, 2H, H2^B, H4^B), 5.02 (d, $J_{1B,2B}$ = 7.8 Hz, 1H, H1^B), 4.85 (d, $J_{1A,2A}$ = 1.1 Hz, 1H, H1^A), 4.63 (dd, $J_{6B,6B'}$ = 12.1 Hz, $J_{6B,5B}$ = 3.1 Hz, 1H, H6^B), 4.50 (dd, $J_{6B',6B}$ = 12.1 Hz, $J_{6B',5B}$ = 6.2 Hz, 1H, H6$^{B'}$), 4.18 (ddd, $J_{5B,4B}$ = 9.5 Hz, $J_{5B,6B'}$ = 6.2 Hz, $J_{5B,6B}$ = 3.1 Hz, 1H, H5^B), 3.90 (dd, $J_{2A,3A}$ = 3.6 Hz, $J_{2A,1A}$ = 1.4 Hz, 1H, H2^A), 3.66 (dd, $J_{3A,4A}$ = 9.5 Hz, $J_{3A,2A}$ = 3.5 Hz, 1H, H3^A), 3.52 (dd, $J_{5A,4A}$ = 9.3 Hz, $J_{5A,6A}$ = 6.2 Hz, 1H, H5^A), 3.25 – 3.23 (m, 1H, H4^A), 3.20 (s, 3H, OCH_3), 1.26 (d, $J_{6A,5A}$ = 6.2 Hz, 3H, H6^A). **APT NMR** (101 MHz, $CDCl_3$) δ 166.2 (1C, C_q), 165.9 (1C, C_q), 165.5 (1C, C_q), 165.3 (1C, C_q) (CO, Bz), 133.8 - 128.5 (Ar*C*H, Ar*C*), 103.3 (C1^B), 99.9 (C1^A), 80.8 (C2^A), 73.9 (C4^A), 72.8 (C3^B), 72.6, 72.5 (C2^B, C5^B), 71.4 (C3^A), 69.7 (C4^B), 67.6 (C5^B), 63.3 (C6^B), 54.8 (O*C*H$_3$), 17.5 (C6^A). The spectra contain small impurities and solvent peaks.

Phenyl 2,3,4,6-tetra-*O*-benzoyl-β-D-glucopyranosyl-(1→4)-1-thio-α-L-rhamnopyranoside (209)



^{1}H NMR (400 MHz, $CDCl_3$) δ 8.05 – 8.03 (m, 2H, Ar*H*), 8.00 – 7.98 (m, 2H, Ar*H*), 7.98 – 7.91 (m, 2H, Ar*H*), 7.83 – 7.81 (m, 2H, Ar*H*), 7.57 – 7.24 (m, 12H, Ar*H*), 5.93 (t, J = 9.7 Hz, 1H, H3^B), 5.67 (t, J = 9.7 Hz, 1H, H4^B), 5.55 (dd, $J_{2B,3B}$ = 9.8 Hz, $J_{2B,1B}$ = 8.0 Hz, 1H, H2^B), 5.42 (d, $J_{1A,2A}$ = 1.1 Hz, 1H, H1^A), 5.25 (d, $J_{1B,2B}$ = 7.9 Hz, 1H, H1^B), 4.70 (dd, $J_{6B,6B'}$ = 12.1 Hz, $J_{6B,5B}$ = 3.0 Hz, 1H, H6^B), 4.48 (dd, $J_{6B',6B}$ = 12.2 Hz, $J_{6B',5B}$ = 5.6 Hz, 1H, H6$^{B'}$), 4.21 – 4.16 (m, 2H, H5^A, H5^B), 4.06 (dd, $J_{2A,3A}$ = 3.3 Hz, $J_{2A,1A}$ = 1.5 Hz, 1H, H2^A), 3.80 (dd, $J_{3A,4A}$ = 9.3 Hz, $J_{3A,2A}$ = 3.3 Hz, 1H, H3^A), 3.70 (t, J = 9.4 Hz, 1H, H4^A), 1.25 (d, $J_{6A,5A}$ = 6.2 Hz, 3H, H6^A). **APT NMR** (101 MHz, $CDCl_3$) δ 166.3 (1C, C_q), 165.9 (1C, C_q), 165.4 (1C, C_q), 165.3 (1C, C_q) (CO, Bz), 133.9 - 127.6 (Ar*C*H, Ar*C*), 101.1 (C1^B), 87.3 (C1^A), 81.8 (C4^A), 73.1 (C3^B), 72.7 (C2^A), 72.5, 72.4 (C2^B, C5^B), 71.3 (C3^A), 69.7 (C4^B), 67.5 (C5^A), 63.0 (C6^B), 17.6 (C6^A). NMR data are in accordance with literature values.[220]

Methyl 2,3,4,6-tetra-*O*-benzoyl-β-D-glucopyranosyl-(1→4)-α-L-rhamnopyranoside (210)



^{1}H NMR (400 MHz, $CDCl_3$) δ 8.05 – 8.02 (m, 2H, Ar*H*), 7.98 – 7.96 (m, 2H, Ar*H*), 7.92 – 7.90 (m, 2H, Ar*H*), 7.82 – 7.80 (m, 2H, Ar*H*), 7.58 – 7.28 (m, 12H, Ar*H*), 5.91 (t, J = 9.7 Hz, 1H, H3^{B}), 5.65 (t, J = 9.7 Hz, 1H, H4^{B}), 5.51 (dd, $J_{2B,3B}$ = 9.8 Hz, $J_{2B,1B}$ = 8.0 Hz, 1H, H2^{B}), 5.22 (d, $J_{1B,2B}$ = 8.0 Hz, 1H, H1^{B}), 4.68 (dd, $J_{6B,6B'}$ = 12.1 Hz, $J_{6B,5B}$ = 3.0 Hz, 1H, H6^{B}), 4.59 (d, $J_{1A,2A}$ = 1.4 Hz, 1H, H1^{A}), 4.47 (dd, $J_{6B',6B}$ = 12.1 Hz, $J_{6B',5B}$ = 5.6 Hz, 1H, H6^{B}'), 4.19 – 4.14 (m, 1H, H5^{B}), 3.78 (dd, $J_{2A,3A}$ = 3.5 Hz, $J_{2A,1A}$ = 1.6 Hz, 1H, H2^{A}), 3.74 (dd, $J_{3A,4A}$ = 8.8 Hz, $J_{3A,2A}$ = 3.5 Hz, 1H, H3^{A}), 3.67 – 3.56 (m, 2H, H4^{A}, H5^{A}), 3.30 (s, 3H, OCH_3), 1.26 (d, $J_{6A,5A}$ = 5.9 Hz, 3H, H6^{A}). **^{13}C NMR** (101 MHz, $CDCl_3$) δ 166.2 (1C, C_q), 165.9 (1C, C_q), 165.4 (1C, C_q), 165.3 (1C, C_q) (CO, Bz), 133.6 - 128.4 (Ar*C*H, Ar*C*), 101.2 (C1^{B}), 100.3 (C1^{A}), 81.6 (C4^{A}), 73.1 (C3^{B}), 72.4, 72.4 (C2^{B}, C5^{B}), 71.2, 70.0 (C2^{A}, C3^{A}), 69.8 (C4^{B}), 65.9 (C5^{A}), 63.1 (C6^{B}), 55.0 (O*C*H$_3$), 17.7 (C6^{A}). The spectra contain small impurities and are not dry. NMR data are in accordance with literature values.[226]

6. Conclusion

This thesis emphasises the strength of chemical synthesis for the formation of xylan fragments, which can be used for elucidation of the specificity of both known and novel xylan degrading enzymes. A general protocol for the synthesis of both arabinoxylan and glucuronoxylan fragments has been utilised to afford five different fragments. The protocol made the synthesis of especially the arabinoxylan fragments possible in higher amounts than the automated synthesis reported by Pfrengle and co-workers for similar compounds.[73,77] Moreover, the overall synthetic strategy could in the future with minor changes be adapted to accommodate the synthesis of other biochemically relevant fragments, where an example could be to install *S*-linked substituents, which could be used as putative inhibitors of xylan degrading enzymes.

Ferulic acid esterases are important accessory enzymes for the enzymes needed for degradation of lignocellulose. This thesis presents the characterisation of 11 ferulic acid esterases with regard to their reactivity and specificity. The results showed the validation of a previous classification of the enzymes into 13 subfamilies based on genome sequences for at least six subfamilies. Thereby, the subfamily classification can hopefully be applied to predict the reactivity of novel ferulic acid esterases.

Regioselective glycosylations are still an important field in carbohydrate chemistry, which has the ability to simplify and shorten the synthesis of oligosaccharides. In this thesis, it was found

that several molybdenum complexes are not able to catalyse regioselective glycosylations, and future studies should therefore focus on other metals or at least other systems using these catalysts.

7. Bibliography

(1) Zhou, C.-H.; Xia, X.; Lin, C.-X.; Tong, D.-S.; Beltramini, J. *Chem. Soc. Rev.* **2011**, *40*, 5588–5617.

(2) Mood, S. H.; Golfeshan, A. H.; Tabatabaei, M.; Jouzani, G. S.; Najafi, G. H.; Gholami, M.; Ardjmand, M. *Renew. Sustain. Energy Rev.* **2013**, *27*, 77–93.

(3) Claassen, P. A. M.; van Lier, J. B.; Contreras, A. M. L.; van Niel, E. W. J.; Sijtsma, L.; Stams, A. J. M.; de Vries, S. S.; Weusthuis, R. A. *Appl. Microbiol. Biotechnol.* **1999**, *52*, 741–755.

(4) Jordan, D. B.; Bowman, M. J.; Braker, J. D.; Dien, B. S.; Hector, R. E.; Lee, C. C.; Mertens, J. A.; Wagschal, K. *Biochem. J.* **2012**, *442*, 241–252.

(5) Peng, F.; Peng, P.; Xu, F.; Sun, R. C. *Biotechnol. Adv.* **2012**, *30*, 879–903.

(6) Zhao, H.; Jones, C. L.; Baker, G. A.; Xia, S.; Olubajo, O.; Person, V. N. *J. Biotechnol.* **2009**, *139*, 47–54.

(7) Alvira, P.; Tomás-Pejó, E.; Ballesteros, M.; Negro, M. J. *Bioresour. Technol.* **2010**, *101*, 4851–4861.

(8) Sun, S.; Sun, S.; Cao, X.; Sun, R. *Bioresour. Technol.* **2016**, *199*, 49–58.

(9) Gibson, G. R.; Probert, H. M.; Loo, J. V.; Rastall, R. A.; Roberfroid, M. B. *Nutr. Res. Rev.* **2004**, *17*, 259–275.

(10) Lindberg, J. E. *J. Anim. Sci. Biotech.* **2014**, *5*, 15–22.

(11) Jha, R.; Berrocoso, J. D. *Animal* **2015**, *9*, 1441–1452.

(12) Scheller, H. V.; Ulvskov, P. *Annu. Rev. Plant Biol.* **2010**, *61*, 263–289.

(13) Ebringerova, A.; Hromadkova, Z.; Heinze, T. *Adv. Polym. Sci.* **2005**, *186*, 1–67.

(14) Li, X.; Chapple, C. *Plant Physiol.* **2010**, *154*, 449–452.

(15) Song, Q.; Wang, F.; Cai, J.; Wang, Y.; Zhang, J.; Yu, W.; Xu, J. *Energ. Environ. Sci.* **2013**, *6*, 994–1007.

(16) Takahashi, N.; Koshijima, T. *Wood Sci. Technol.* **1988**, *22*, 231–241.

(17) Lam, T. B. T.; Kadoya, K.; Iiyama, K. *Phytochemistry* **2001**, *57*, 987–992.

(18) Yapo, B. M. *Carbohydr. Polym.* **2011**, *86*, 373–385.

(19) Harholt, J.; Suttangkakul, A.; Scheller, H. V. *Physiol. Plant.* **2007**, *129*, 283–295.

(20) Mendis, M.; Leclerc, E.; Simsek, S. *Carbohydr. Polym.* **2016**, *139*, 159–166.

(21) Sjöström, E. In *Wood Chemistry, Fundamentals and Application*; Academic Press, Inc., 1993; pp 51–70.

(22) Teleman, A.; Siika-Aho, M.; Sorsa, H.; Buchert, J.; Perttula, M.; Hausalo, T.; Tenkanen, M. *Carbohydr. Res.* **1996**, *293*, 1–13.

(23) Rydlund, A.; Dahlman, O. *Carbohydr. Res.* **1997**, *300*, 95–102.

(24) Balakshin, M.; Capanema, E.; Berlin, A. In *Studies in Natural Products Chemistry*; Elsevier B.V., 2014; pp 83–115.

(25) Buanafina, M. M. de O. *Mol. Plant* **2009**, *2*, 861–872.

(26) Wong, D. W. S. *Appl. Biochem. Biotechnol.* **2006**, *133*, 87–112.

(27) Dilokpimol, A.; Mäkelä, M. R.; Aguilar-Pontes, M. V.; Benoit-Gelber, I.; Hildén, K. S.; de Vries, R. P. *Biotechnol. Biofuels* **2016**, *9*, 231–240.

(28) Ou, S.; Kwok, K.-C. *J. Sci. Food Agric.* **2004**, *84*, 1261–1269.

(29) Ralet, M. C.; Thibault, J. F.; Faulds, C. B.; Williamson, G. *Carbohydr. Res.* **1994**, *263*, 227–241.

(30) Waldron, K. W.; Parr, A. J.; Ng, A.; Ralph, J. *Phytochem. Anal.* **1996**, *7*, 305–312.

(31) Vismeh, R.; Lu, F.; Chundawat, S. P. S.; Humpula, J. F.; Azarpira, A.; Balan, V.; Dale, B. E.; Ralph, J.; Jones, A. D. *Analyst* **2013**, *138*, 6683–6692.

(32) Ralph, J. *Phytochem. Rev.* **2010**, *9*, 65–83.

(33) de Oliveira, D. M.; Finger-Teixeira, A.; Mota, T. R.; Salvador, V. H.; Moreira-Vilar, F. C.; Molinari, H. B. C.; Mitchell, R. A. C.; Marchiosi, R.; Ferrarese-Filho, O.; dos Santos, W. D. *Plant Biotechnol. J.* **2015**, *13*, 1224–1232.

(34) Rouau, X.; Cheynier, V.; Surget, A.; Gloux, D.; Barron, C.; Meudec, E.; Louis-Montero, J.; Criton, M. *Phytochemistry* **2003**, *63*, 899–903.

(35) Funk, C.; Ralph, J.; Steinhart, H.; Bunzel, M. *Phytochemistry* **2005**, *66*, 363–371.

(36) Bunzel, M.; Ralph, J.; Funk, C.; Steinhart, H. *Tetrahedron Lett.* **2005**, *46*, 5845–5850.

(37) Dodd, D.; Cann, I. O. *Glob. Chang. Biol. Bioenergy* **2009**, *1*, 2–17.

(38) Monclaro, A. V.; Filho, E. X. F. In *Fungal Enzymes*; CRC Press, Taylor & Francis Group, 2014; pp 280–293.

(39) Gottschalk, L. M. F.; Oliveira, R. A.; Bon, E. P. da S. *Biochem. Eng. J.* **2010**, *51*, 72–78.

(40) de Vries, R. P.; Kester, H. C. M.; Poulsen, C. H.; Benen, J. A. E.; Visser, J. *Carbohydr. Res.* **2000**, *327*, 401–410.

(41) Koshland, D. E. *Biol. Rev.* **1953**, *28*, 416–436.

(42) Biely, P.; Singh, S.; Puchart, V. *Biotechnol. Adv.* **2016**, *34*, 1260–1274.

(43) Lagaert, S.; Pollet, A.; Courtin, C. M.; Volckaert, G. *Biotechnol. Adv.* **2014**, *32*, 316–332.

(44) Kolenová, K.; Vršanská, M.; Biely, P. *J. Biotechnol.* **2006**, *121*, 338–345.

(45) Maslen, S. L.; Goubet, F.; Adam, A.; Dupree, P.; Stephens, E. *Carbohydr. Res.* **2007**, *342*, 724–735.

(46) Firbank, S. J.; Zhu, Y.; Lewis, R. J.; York, W. S.; Fontes, C. M. G. A.; Gilbert, H. J. *J. Biol. Chem.* **2011**, *286*, 22510–22520.

(47) Correia, M. A. S.; Mazumder, K.; Brás, J. L. A.; Firbank, S. J.; Zhu, Y.; Lewis, R. J.; York, W. S.; Fontes, C. M. G. A.; Gilbert, H. J. *J. Biol. Chem.* **2011**, *286*, 22510–22520.

(48) Nishitani, K.; Nevins, D. J. *J. Biol. Chem.* **1991**, *266*, 6539–6543.

(49) Biely, P.; Puchart, V.; Stringer, M. A.; Krogh, K. B. R. M. *FEBS J.* **2014**, *281*, 3894–3903.

(50) Sørensen, H. R.; Jørgensen, C. T.; Pedersen, S.; Meyer, A. S. *Appl. Microbiol. Biotechnol.* **2006**, *73*, 850–861.

(51) Puls, J.; Schmidt, O.; Granzow, C. *Enzym. Microb. Technol.* **1987**, *9*, 83–88.

(52) Ryabova, O.; Vršanská, M.; Kaneko, S.; van Zyl, W. H.; Biely, P. *FEBS Lett.* **2009**, *583*, 1457–1462.

(53) Chong, S. L.; Battaglia, E.; Coutinho, P. M.; Henrissat, B.; Tenkanen, M.; de Vries, R. P.

Appl. Microbiol. Biotechnol. **2011**, *90*, 1323–1332.

(54) Faulds, C. B. *Phytochem. Rev.* **2010**, *9*, 121–132.

(55) Mackenzie, C. R.; Bilous, D.; Schneider, H.; Johnson, K. G. *Appl. Environ. Microbio.* **1987**, *53*, 2835–2839.

(56) Mackenzie, C. R.; Bilous, D. *Appl. Environ. Microbiol.* **1988**, *54*, 1170–1173.

(57) Johnson, K. G.; Harrison, B. A.; Schneider, H.; MacKenzie, C. R.; Fontana, J. D. *Enzym. Microb. Technol.* **1988**, *10*, 403–409.

(58) Topakas, E.; Vafiadi, C.; Christakopoulos, P. *Process Biochem.* **2007**, *42*, 497–509.

(59) Katsimpouras, C.; Antonopoulou, I.; Christakopoulos, P.; Topakas, E. In *Microbial Enzymes in Bioconversions of Biomass*; Springer, 2016; pp 79–125.

(60) Kraut, J. *Annu. Rev. Biochem.* **1977**, *46*, 331–358.

(61) Prates, J. A. M.; Tarbouriech, N.; Charnock, S. J.; Fontes, C. M. G. A.; Ferreira, L. M. A.; Davies, G. J. *Structure* **2001**, *9*, 1183–1190.

(62) Berg, J. M.; Tymoczko, J. L.; Stryer, L. In *Biochemistry*; W. H. Freeman and Company, 2007; pp 241–274.

(63) Crepin, V. F.; Faulds, C. B.; Connerton, I. F. *Appl. Biochem. Biotechnol.* **2004**, *63*, 647–652.

(64) Benoit, I.; Danchin, E. G. J.; Bleichrodt, R. J.; de Vries, R. P. *Biotechnol. Lett.* **2008**, *30*, 387–396.

(65) Udatha, D. B. R. K. G.; Kouskoumvekaki, I.; Olsson, L.; Panagiotou, G. *Biotechnol. Adv.* **2011**, *29*, 94–110.

(66) Dilokpimol, A.; Mäkelä, M. R.; Varriale, S.; Zhou, M.; Cerullo, G.; Gidijala, L.; Hinkka, H.; Brás, J. L. A.; Jütten, P.; Piechot, A.; et al. *N. Biotechnol.* **2018**, *41*, 9–14.

(67) Walton, N. J.; Narbad, A.; Faulds, C. B.; Williamson, G. *Curr. Opin. Biotechnol.* **2000**, *11*, 490–496.

(68) Tian, L.; Gruppen, H.; Schols, H. A. *J. Agric. Food Chem.* **2015**, *63*, 10822–10830.

(69) Petersen, B. O.; Lok, F.; Meier, S. *Carbohydr. Polym.* **2014**, *112*, 587–594.

(70) Hirsch, J.; Kováč, P.; Petráková, E. *Carbohydr. Res.* **1982**, *106*, 203–216.

(71) Kováč, P.; Hirsch, J. *Carbohydr. Res.* **1982**, *100*, 177–193.

(72) Takeo, K.; Ohguchi, Y.; Hasegawa, R.; Kitamura, S. *Carbohydr. Res.* **1995**, *278*, 301–313.

(73) Schmidt, D.; Schuhmacher, F.; Geissner, A.; Seeberger, P. H.; Pfrengle, F. *Chem. Eur. J.* **2015**, *21*, 5709–5713.

(74) Hirsch, J.; Petráková, E. *Carbohydr. Res.* **1984**, *131*, 219–226.

(75) Kováč, P. *J. Carbohydr. Nucleos. Nucleot.* **1977**, *4*, 165–173.

(76) Oscarson, S.; Svahnberg, P. *J. Chem. Soc. Perkin Trans. 1* **2001**, 873–879.

(77) Senf, D.; Ruprecht, C.; Kruijff, G. H. M. De; Simonetti, S. O.; Schuhmacher, F.; Seeberger, P. H.; Pfrengle, F. *Eur. J. Chem.* **2017**, *23*, 3197–3205.

(78) Kováč, P.; Hirsch, J.; Kováčik, V.; Kočiš, P. *Carbohydr. Res.* **1980**, *85*, 41–49.

(79) Myhre, D. V.; Smith, F. *J. Org. Chem.* **1961**, *26*, 4609–4612.

(80) Helferich, B.; Wedemeyer, K.-F. *Ann. Chem.* **1948**, *990*, 139–145.

(81) Koenigs, W.; Knorr, E. *Chem. Ber.* **1901**, *34*, 957–981.

(82) Kováč, P.; Hirsch, J. *Carbohydr. Res.* **1981**, *90*, C5–C7.

(83) Ferrier, R. J.; Hay, R. W.; Vethaviyasar, N. *Carbohydr. Res.* **1973**, *27*, 55–61.

(84) Zhong, W.; Boons, G.-J. In *Handbook of Chemical Glycosylation*; Demchenko, A. V., Ed.; Wiley VCH, 2008; pp 261–357.

(85) Veeneman, G. H.; van Leeuwen, S. H.; van Boom, J. H. *Tetrahedron Lett.* **1990**, *31*, 1331–1334.

(86) Fügedi, P.; Garegg, P. J. *Carbohydr. Res.* **1986**, *149*, C9–C12.

(87) Huang, X.; Huang, L.; Wang, H.; Ye, X. *Angew. Chem. Int. Ed.* **2004**, *43*, 5221–5224.

(88) Crich, D.; Smith, M. *J. Am. Chem. Soc.* **2001**, *123*, 9015–9020.

(89) Codée, J. D. C.; Litjens, R. E. J. N.; den Heeten, R.; Overkleeft, H. S.; van Boom, J. H.; van der Marel, G. A. *Org. Lett.* **2003**, *5*, 1519–1522.

(90) Martichonok, V.; Whitesides, G. M. *J. Org. Chem.* **1996**, *61*, 1702–1706.

(91) Crich, D.; Cai, F.; Yang, F. *Carbohydr. Res.* **2008**, *343*, 1858–1862.

(92) Lian, G.; Zhang, X.; Yu, B. *Carbohydr. Res.* **2015**, *403*, 13–22.

(93) Gao, J.; Guo, Z. *J. Org. Chem.* **2013**, *78*, 12717–12725.

(94) Kováč, P.; Hirsch, J.; Kováčik, V. *Carbohydr. Res.* **1979**, *75*, 109–116.

(95) Fürst, A.; Plattner, P. A. *Helv. Chim. Acta.* **1949**, *32*, 275–283.

(96) Kováč, P.; Hirsch, J. *Carbohydr. Res.* **1980**, *79*, 303–307.

(97) Kováč, P.; Hirsch, J.; Kováčik, V. *Chem. Zvesti.* **1978**, *32*, 514–518.

(98) van den Bos, L. J.; Codée, J. D. C.; Litjens, R. E. J. N.; Dinkelaar, J.; Overkleeft, H. S.; van der Marel, G. A. *Eur. J. Org. Chem.* **2007**, 3963–3976.

(99) Codée, J. D. C.; Christina, A. E.; Walvoort, M. T. C.; Overkleeft, H. S.; van der Marel, G. A. *Top. Curr. Chem.* **2011**, *301*, 253–289.

(100) Edward, J. T. *Chem. Ind.* **1955**, 1102–1104.

(101) Kuszmann, J. In *The Organic Chemistry of Sugars*; Taylor & Francis Group, 2006; pp 25–52.

(102) Cumpstey, I. *Org. Biomol. Chem.* **2012**, *10*, 2503–2508.

(103) Mong, K.-K. T.; Nokami, T.; Tran, N. T. T.; Nhi, P. B. In *Selective Glycosylations: Synthetic Methods and Catalysts*; Wiley-VCH, 2017; pp 59–77.

(104) Alpe, M.; Oscarson, S. *Carbohydr. Res.* **2003**, *338*, 2605–2609.

(105) Böhm, M. F. *Ph.D. Thesis: Chemical Synthesis of Hemicellulose Fragments*, Technical University of Denmark, 2016.

(106) Tatai, J.; Fügedi, P. *Org. Lett.* **2007**, *9*, 4647–4650.

(107) Huang, T.-Y.; Zulueta, M. M. L.; Hung, S.-C. *Org. Lett.* **2011**, *13*, 1506–1509.

(108) Crich, D.; Smith, M.; Yao, Q.; Picione, J. *Synthesis (Stuttg).* **2001**, 323–326.

(109) Nicolaou, K. C.; Rodríguez, R. M.; Fylaktakidou, K. C.; Suzuki, H.; Mitchell, H. J. *Angew. Chem. Int. Ed.* **1999**, *38*, 3340–3345.

(110) López, R.; Fernández-Mayoralas, A. *J. Org. Chem.* **1994**, *59*, 737–745.

(111) Kim, J. M.; Roy, R. *Carbohydr. Res.* **1997**, *298*, 173–179.

(112) Oikawa, Y.; Yoshioka, T.; Yonemitsu, O. *Tetrahedron Lett.* **1982**, *23*, 885–888.

(113) Thorsheim, K.; Siegbahn, A.; Johnsson, R. E.; Stålbrand, H.; Manner, S.; Widmalm, G.; Ellervik, U. *Carbohydr. Res.* **2015**, *418*, 65–88.

(114) Rio, S.; Beau, J. M.; Jacquinet, J. C. *Carbohydr. Res.* **1991**, *219*, 71–90.

(115) Yang, B.; Yoshida, K.; Yin, Z.; Dai, H.; Kavunja, H.; El-Dakdouki, M. H.; Sungsuwan, S.; Dulaney, S. B.; Huang, X. *Angew. Chem. Int. Ed.* **2012**, *51*, 10185–10189.

(116) Andersen, S. M.; Heuckendorff, M.; Jensen, H. H. *Org. Lett.* **2015**, *17*, 944–947.

(117) Tully, S. E.; Mabon, R.; Gama, C. I.; Tsai, S. M.; Liu, X.; Hsieh-Wilson, L. C. *J. Am. Chem. Soc.* **2004**, *126*, 7736–7737.

(118) Kocienski, P. J. In *Protection Groups*; Georg Thieme Verlag, 1994; pp 1–16.

(119) Garegg, P. J.; Kvarnström, I.; Niklasson, A.; Niklasson, G.; Svensson, S. C. T. *J. Carbohydr. Chem.* **1993**, *12*, 933–953.

(120) Haines, B. Y. A. H. *Adv. Carbohydr. Chem. Biochem.* **1976**, *33*, 11–109.

(121) Roslund, M. U.; Aitio, O.; Wä, J.; Maaheimo, H.; Murzin, D. Y.; Leino, R. *J. Am. Chem. Soc.* **2008**, *13*, 8769–8772.

(122) Daugaard, M. *Ph.D. Thesis: Synthesis of Oligo (1→5)-α-L-Arabinofuranosides Related to the Plant Polysaccharide Pectin*, Technical University of Denmark, 2016.

(123) Callam, C. S.; Lowary, T. L. *J. Chem. Educ.* **2001**, *78*, 73–74.

(124) Dai, Y.; Yu, B. *Chem. Commun.* **2015**, *51*, 13826–13829.

(125) Andersen, M. C. F.; Kracun, S. K.; Rydahl, M. G.; Willats, W. G. T.; Clausen, M. H. *Chem. Eur. J.* **2016**, *22*, 11543–11548.

(126) Crich, D.; Sun, S. *J. Am. Chem. Soc.* **1997**, *119*, 11217–11223.

(127) Crich, D.; Cai, W. *J. Org. Chem.* **1999**, *64*, 4926–4930.

(128) Frihed, T. G.; Bols, M.; Pedersen, C. M. *Chem. Rev.* **2015**, *115*, 4963–5013.

(129) Crich, D.; Chandrasekera, N. S. *Angew. Chem. Int. Ed.* **2004**, *43*, 5386–5389.

(130) Bohé, L.; Crich, D. *Carbohydr. Res.* **2015**, *403*, 48–59.

(131) Martin, A.; Arda, A.; Désiré, J.; Martin-Mingot, A.; Probst, N.; Sinaÿ, P.; Jiménez-Barbero, J.; Thibaudeau, S.; Blériot, Y. *Nat. Chem.* **2016**, *8*, 186–191.

(132) Geurtsen, R.; Boons, G. J. *Tetrahedron Lett.* **2002**, *43*, 9429–9431.

(133) Li, Z.; Gildersleeve, J. C. *J. Am. Chem. Soc.* **2006**, *128*, 11612–11619.

(134) Christensen, H. M.; Oscarson, S.; Jensen, H. H. *Carbohydr. Res.* **2015**, *408*, 51–95.

(135) Glushka, J. N.; Perlin, A. S. *Carbohydr. Res.* **1990**, *205*, 305–321.

(136) Bock, K.; Pedersen, C. *J. Chem. Soc. Perkin Trans. 2* **1974**, 293–297.

(137) Mizutani, K.; Kasai, R.; Nakamura, M.; Tanaka, O.; Matsuura, H. *Carbohydr. Res.* **1989**, *185*, 27–38.

(138) Gellerstedt, G. In *Pulping and Paper Chemistry and Technology*; de Gruyter GmbH, 2009; pp 92–119.

(139) Wang, Y.; Azhar, S.; Lindström, M. E.; Henriksson, G. *J. Wood Chem. Technol.* **2015**, *35*, 91–101.

(140) Martínez, P. M. *Ph.D. Thesis: Alkaline Pretreatment of Lignin-Rich by-Products and Their Implications for Enzymatic Degradation*, Wageningen University and Research, 2016.

(141) Oba, M.; Tanaka, K.; Nishiyama, K.; Ando, W. *J. Org. Chem.* **2011**, *76*, 4173–4177.

(142) Das, R.; Chakraborty, D. *Tetrahedron Lett.* **2010**, *51*, 6255–6258.

(143) Tian, X. X.; Li, A.; Farrugia, I. V.; Mo, X.; Crich, D.; Groves, M. J. *Carbohydr. Res.* **2000**, *324*, 38–44.

(144) Williams, D. B. G.; Lawton, M. *J. Org. Chem.* **2010**, *75*, 8351–8354.

(145) Kongkathip, B.; Kongkathip, N.; Rujirawanic, J. *Synth. Commun.* **2014**, *44*, 2248–2255.

(146) Crich, D.; Dai, Z. *Tetrahedron* **1999**, *55*, 1569–1580.

(147) Phanumartwiwath, A.; Hornsby, T. W.; Jamalis, J.; Bailey, C. D.; Willis, C. L. *Org. Lett.* **2013**, *15*, 5734–5737.

(148) Bonora, B.; Boos, I.; Clausen, M. H. *Carbohydr. Res.* **2017**, *443–444*, 53–57.

(149) van Rijssel, E. R.; Goumans, T. P. M.; Lodder, G.; Overkleeft, H. S.; van der Marel, G. A.; Codée, J. D. C. *Org. Lett.* **2013**, *15*, 3026–3029.

(150) D'Errico, C. *Ph.D. Thesis: Glucuronic Acid Derivatives in Enzymatic Biomass Degradation: Synthesis and Evaluation of Enzymatic Activity*, Technical University of Denmark, 2016.

(151) Rej, R. N.; Glushka, J. N.; Chew, W.; Perlin, A. S. *Carbohydr. Res.* **1989**, *189*, 135–148.

(152) Nicolaou, K. C.; Mitchell, H. J.; Fylaktakidou, K. C.; Rodríguez, R. M.; Suzuki, H. *Chem. Eur. J.* **2000**, *6*, 3116–3148.

(153) Wutz, P. G. M.; Greene, T. W. In *Greene's protective groups in organic synthesis*; John Wiley and Sons, Inc., 2007.

(154) Ogilvie, K. K.; Entwistle, D. W. *Carbohydr. Res.* **1981**, *89*, 203–210.

(155) Jones, S. S.; Reese, C. B. *J. Chem. Soc. Perkin Trans. 1* **1979**, 2762–2764.

(156) Clayden, J.; Greeves, N.; Warren, S.; Wothers, P. In *Organic Chemistry*; 2001; pp 1246–1310.

(157) Garegg, P. J.; Olsson, L.; Oscarson, S. *J. Org. Chem.* **1995**, *60*, 2200–2204.

(158) Houdier, S.; Vottero, P. J. A. *Carbohydr. Res.* **1993**, *248*, 377–384.

(159) van den Bos, L. J.; Codée, J. D. C.; van der Toorn, J. C.; Boltje, T. J.; van Boom, J. H.; Overkleeft, H. S.; van der Marel, G. A. *Org. Lett.* **2004**, *6*, 2165–2168.

(160) Bera, S.; Linhardt, R. J. *J. Org. Chem.* **2011**, *76* (9), 3181–3193.

(161) Shih, T. L.; Fang, Y. C. *Synth. Commun.* **2007**, *37*, 3337–3349.

(162) Crich, D.; Vinogradova, O. *J. Org. Chem.* **2007**, *72*, 3581–3584.

(163) Debenham, J. S.; Rodebaugh, R.; Fraser-Reid, B. *J. Org. Chem.* **1997**, *62*, 4591–4600.

(164) Lemieux, R. U.; Hendriks, K. B.; Stick, R. V.; James, K. *J. Am. Chem. Soc.* **1975**, *97*, 4056–4062.

(165) Pedersen, C. M.; Nordstrøm, L. U.; Bols, M. *J. Am. Chem. Soc.* **2007**, *129*, 9222–9235.

(166) Lucas, H.; Basten, J. E. M.; van Dinther, T. G.; Meuleman, D. G.; van Aelst, S. F.; van

Boeckel, C. A. A. *Tetrahedron* **1990**, *46*, 8207–8228.

(167) Jacquinet, J. C.; Lopin-Bon, C.; Vibert, A. *Chem. Eur. J.* **2009**, *15*, 9579–9595.

(168) Codée, J. D. C.; van den Bos, L. J.; de Jong, A. R.; Dinkelaar, J.; Lodder, G.; Overkleeft, H. S.; van der Marel, G. A. *J. Org. Chem.* **2009**, *74*, 38–47.

(169) Zhang, Q.; Van Rijssel, E. R.; Walvoort, M. T. C.; Overkleeft, H. S.; Van Der Marel, G. A.; Codée, J. D. C. *Angew. Chem. Int. Ed.* **2015**, *54*, 7670–7673.

(170) Dhamale, O. P.; Zong, C.; Al-Mafraji, K.; Boons, G. J. *Org. Biomol. Chem.* **2014**, *12*, 2087–2098.

(171) Salamone, S.; Boisbrun, M.; Didierjean, C.; Chapleur, Y. *Carbohydr. Res.* **2014**, *386*, 99–105.

(172) Takahashi, H.; Shida, T.; Hitomi, Y.; Iwai, Y.; Miyama, N.; Nishiyama, K.; Sawada, D.; Ikegami, S. *Chem. Eur. J.* **2006**, *12*, 5868–5877.

(173) Takahashi, H.; Miyama, N.; Mitsuzuka, H.; Ikegami, S. *Synthesis (Stuttg).* **2004**, *18*, 2991–2994.

(174) van den Bos, L. J.; Codée, J. D. C.; van der Toorn, J. C.; Boltje, T. J.; van Boom, J. H.; Overkleeft, H. S.; van der Marel, G. A. *Org. Lett.* **2004**, *6*, 2165–2168.

(175) Verduyn, R.; Douwes, M.; van der Klein, P. A. M.; Mösinger, E. M.; van der Marel, G. A.; van Boom, J. H. *Tetrahedron* **1993**, *49*, 7301–7316.

(176) Kumar, A.; Geng, Y.; Schmidt, R. R. *Eur. J. Org. Chem.* **2012**, 6846–6851.

(177) Garegg, P. J.; Kvarnström, I.; Niklasson, A.; Niklasson, G.; Svensson, S. C. T. *J. Carbohydr. Chem.* **1993**, *12*, 933–953.

(178) Michihata, N.; Kaneko, Y.; Kasai, T.; Tanigwa, K.; Hirokane, T.; Higasa, S.; Yamada, H. *J. Org. Chem.* **2013**, *78*, 4319–4328.

(179) Kühnel, S.; Pouvreau, L.; Appeldoorn, M. M.; Hinz, S. W. A.; Schols, H. A.; Gruppen, H. *Enzym. Microb. Technol.* **2012**, *50*, 77–85.

(180) Antonopoulou, I.; Dilokpimol, A.; Iancu, L.; Mäkelä, M. R.; Varriale, S.; Cerullo, G.; Hüttner, S.; Uthoff, S.; Jütten, P.; Piechot, A.; et al. *Catalysts* **2018**, *8*, 242–264.

(181) Kálmán, G.; Recseg, K.; Gáspár, M.; Réczey, K. *Appl. Biochem. Biotechnol.* **2006**, *131*, 738–750.

(182) Appeldoorn, M. M.; Kabel, M. A.; van Eylen, D.; Gruppen, H.; Schols, H. A. *J. Agric. Food Chem.* **2010**, *58*, 11294–11301.

(183) Appeldoorn, M. M.; de Waard, P.; Kabel, M. A.; Gruppen, H.; Schols, H. A. *Carbohydr. Res.* **2013**, *381*, 33–42.

(184) Harris, P. J.; Trethewey, J. A. K. *Phytochem. Rev.* **2010**, *9*, 19–33.

(185) Templeton, D. W.; Sluiter, A. D.; Hayward, T. K.; Hames, B. R.; Thomas, S. R. *Cellulose* **2009**, *16*, 621–639.

(186) Berchem, T.; Roiseux, O.; Vanderghem, C.; Boisdnghien, A.; Foucart, G.; Richel, A. *Biofuels, Bioprod. Bioref.* **2017**, *11*, 430–440.

(187) van Dongen, F. E. M.; van Eylen, D.; Kabel, M. A. *Carbohydr. Polym.* **2011**, *86*, 722–731.

(188) van Erven, G.; de Visser, R.; Merkx, D. W. H.; Strolenberg, W.; de Gijsel, P.; Gruppen, H.; Kabel, M. A. *Anal. Chem.* **2017**, *89*, 10907–10916.

(189) Bakker, R.; Elbersen, W.; Poppens, R.; Lesschen, J. P.; Boschma, S.; Kwant, K. W. In *Rice straw and Wheat straw. Potential feedstocks for the Biobased Economy Potential feedstocks for the Biobased Economy, June 2013*; NL Agency Ministry of Economic Affairs, 2013; pp 1–31.

(190) Thermo Fisher Scientific. User Guide: Pierce BCA Protein Assay Kit https://tools.thermofisher.com/content/sfs/manuals/MAN0011430_Pierce_BCA_Protein_Asy_UG.pdf.

(191) Thermo Fisher Scientific. User Guide: Pierce BSA Protein Assay Kit – Reducing Agent Compatible https://www.thermofisher.com/order/catalog/product/23250.

(192) Dilokpimol, A.; Mäkelä, M. R.; Mansouri, S.; Belova, O.; Waterstraat, M.; Bunzel, M.; de Vries, R. P.; Hildén, K. S. *N. Biotechnol.* **2017**, *37*, 200–209.

(193) Faulds, C. B.; Williamson, G. *Microbiology* **1994**, *140*, 779–787.

(194) Harris, D. C. In *Quantitative Chemical Analysis*; W. H. Freeman and Company, 2010.

(195) Bunzel, M.; Ralph, J.; Brüning, P.; Steinhart, H. *J. Agric. Food Chem.* **2006**, *54*, 6409–6418.

(196) Schatz, P. F.; Ralph, J.; Lu, F.; Guzei, I. A.; Bunzel, M. *Org. Biomol. Chem.* **2006**, *4*, 2801–2806.

(197) Humphreys, J. M.; Hemm, M. R.; Chapple, C. *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 10045–10050.

(198) Morreel, K.; Ralph, J.; Lu, F.; Goeminne, G.; Busson, R.; Herdewijn, P.; Goeman, J. L.; van der Eycken, J.; Boerjan, W.; Messens, E. *Plant Physiol.* **2004**, *136*, 4023–4036.

(199) Koseki, T.; Fushinobu, S.; Ardiansyah; Shirakawa, H.; Komai, M. *Appl. Microbiol. Biotechnol.* **2009**, *84*, 803–810.

(200) Reinoso, F. A. M.; Rencoret, J.; Gutiérrez, A.; Milagres, A. M. F.; del Río, J. C.; Ferraz, A. *Biotechnol. Biofuels* **2018**, *11*, 1–12.

(201) Zhu, X.; Schmidt, R. R. *Angew. Chem. Int. Ed.* **2009**, *48*, 1900–1934.

(202) Kanie, O.; Hindsgaul, O. *Curr. Opin. Struct. Biol.* **1992**, *2*, 674–681.

(203) Gyurcsik, B.; Nagy, L. *Coord. Chem. Rev.* **2000**, *203*, 81–149.

(204) Alekseev, Y. E.; Garnovskii, A. D.; Zhdanov, Y. A. *Russ. Chem. Rev.* **1998**, *67*, 649–669.

(205) David, S.; Hanessian, S. *Tetrahedron* **1985**, *41*, 643–663.

(206) Evtushenko, E. V. *Synth. Commun.* **2006**, *36*, 1593–1599.

(207) Evtushenko, E. V. *J. Carbohydr. Chem.* **2010**, *29*, 369–378.

(208) Eby, R.; Schuerch, C. *Carbohydr. Res.* **1982**, *100*, C41–C43.

(209) Grindley, T. B. In *Adv. Carbohydr. Chem. Biochem.*; Academic Press., 1998; pp 17–142.

(210) Ogawa, T.; Matsui, M. *Carbohydr. Res.* **1976**, *51*, C13–C18.

(211) Augé, C.; Veyrières, A. *J. Chem. Soc., Perkin Trans. 1* **1979**, 1825–1832.

(212) Murase, T.; Kartha, K. P. R.; Kiso, M.; Hasegawa, A. *Carbohydr. Res.* **1989**, *195*, 134–137.

(213) Cruzado, C.; Bernabe, M.; Martin-Lomas, M. *Carbohydr. Res.* **1990**, *203*, 296–301.

(214) Garegg, P. J.; Maloisel, J.-L.; Oscarson, S. *Synthesis (Stuttg).* **1995**, 409–414.

(215) Maggi, A.; Madsen, R. *Eur. J. Org. Chem.* **2013**, No. 13, 2683–2691.

(216) Niedbal, D. A.; Madsen, R. *Tetrahedron* **2016**, *72*, 415–419.

(217) Oshima, K.; Aoyama, Y. *J. Am. Chem. Soc.* **1999**, *121*, 2315–2316.

(218) Gouliaras, C.; Lee, D.; Chan, L.; Taylor, M. S. *J. Am. Chem. Soc.* **2011**, *133*, 13926–13929.

(219) Kaji, E.; Nishino, T.; Ishige, K.; Ohya, Y.; Shirai, Y. *Tetrahedron Lett.* **2010**, *51*, 1570–1573.

(220) Fenger, T. H.; Madsen, R. *Eur. J. Org. Chem.* **2013**, 5923–5933.

(221) Lanz, G. *Ph.D. Thesis: Glycosyl Bromides in Glycoside Synthesis: Development of New Promoter System and Metal-Mediated Regioselective Glycosylations*, Technical University of Denmark, 2016.

(222) Sofia, M. J.; Kakarla, R.; Kogan, N.; Dulina, R.; Hui, Y. W.; Hatzenbuhler, N. T.; Liu, D.; Chen, A.; Wagler, T. *Bioorg. Med. Chem. Lett.* **1997**, *7*, 2251–2254.

(223) Groneberg, R. D.; Miyazaki, T.; Stylianides, N. A.; Schulze, T. J.; Stahl, W.; Schreiner, E. P.; Suzuki, T.; Iwabuchi, Y.; Smith, A. L.; Nicolaou, K. C. *J. Am. Chem. Soc.* **1993**, *115*, 7593–7611.

(224) Yuan, X.; Cheng, S.; Shi, Y.; Xue, W. *Synthesis (Stuttg).* **2014**, *46*, 331–335.

(225) Gantt, R. W.; Peltier-Pain, P.; Cournoyer, W. J.; Thorson, J. S. *Nat. Chem. Biol.* **2011**, *7*, 685–691.

(226) Kaji, E.; Nishino, T.; Ishige, K.; Ovya, Y.; Shirai, Y. *Tetrahedron Lett.* **2010**, *51*, 1570–1573.

(227) Pozsgay, V.; Jennings, H. J. *J. Org. Chem.* **1988**, *53*, 4042–4052.

(228) Zhang, G.; Fang, L.; Zhu, L.; Aimiuwu, J. E.; Shen, J.; Cheng, H.; Muller, M. T.; Lee, G. E.; Sun, D.; Wang, P. G. *J. Med. Chem.* **2005**, *48*, 5269–5278.

Appendix

Supporting information for Chapter 4

Table A1: Numerical values for the measured substrate concentrations (mM) and the standard deviations of MF and MpC after 2 and 19 h

Enzyme	SF	MF_{2h} (mM)	Std. $dev._{MF\ 2h}$	MF_{19h} (mM)	Std. $dev._{MF\ 19h}$	MpC_{2h} (mM)	Std. $dev._{MpC\ 2h}$	MpC_{19h} (mM)	Std. $dev._{MpC\ 19h}$
Broth		2.00	0.00	2.00	0.02	2.00	0.04	2.00	0.11
AnFaeB	1	1.80	0.02	1.30	0.11	1.13	0.11	0.07	0.05
AsFaeF	1	1.78	0.03	0.88	0.03	1.65	0.08	0.27	0.01
AnFaeC	5	1.85	0.01	1.64	0.01	1.83	0.06	2.04	0.06
AnidFAEC	5	1.17	0.02	0.01	0.00	0.98	0.09	0.00	0.00
AsFaeC	5	1.59	0.01	0.09	0.00	1.30	0.04	0.03	0.01
C1FaeA1	5	0.03	0.00	0.00	0.00	0.66	0.01	0.00	0.00
C1FaeA2	5	1.00	0.02	0.00	0.00	0.85	0.06	0.00	0.00
FoFae2	6	1.80	0.01	1.27	0.08	1.45	0.06	1.36	0.05
AsFeE	6	1.31	0.04	0.30	0.01	0.92	0.05	0.08	0.01
C1FaeB2	6	1.52	0.00	0.37	0.00	0.88	0.05	0.05	0.00
AnFaeA	7	1.57	0.01	0.50	0.00	2.07	0.16	2.07	0.31
AnFaeJ	9	1.98	0.02	1.69	0.19	2.00	0.12	2.10	0.10
AsFaeI	13	1.50	0.17	0.73	0.02	1.69	0.16	0.22	0.11
ShFae1	13	1.91	0.01	1.91	0.06	1.96	0.02	2.21	0.06

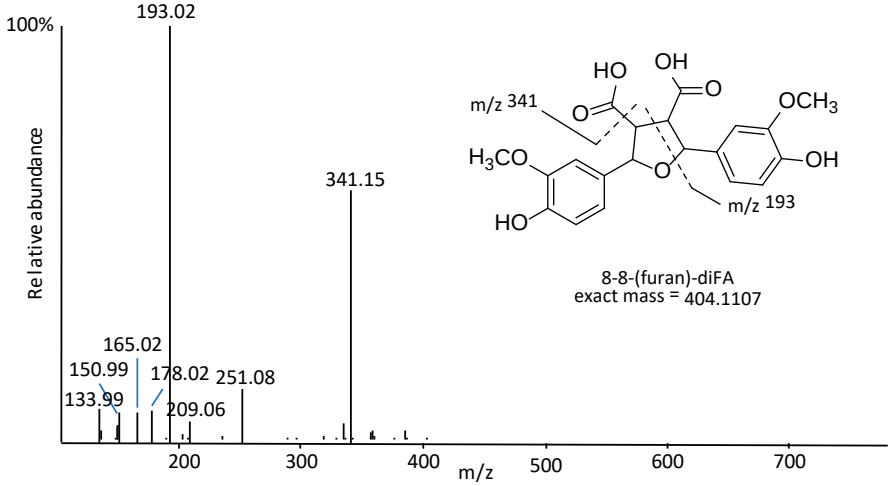


Figure A1: MS² spectrum for 8-8′-(furan)-diFA

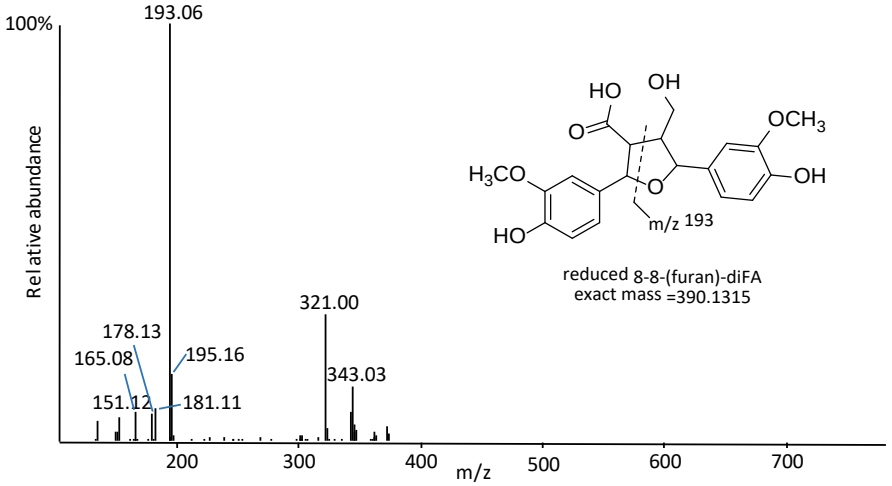


Figure A2: MS² spectrum for reduced 8-8'-(furan)-diFA

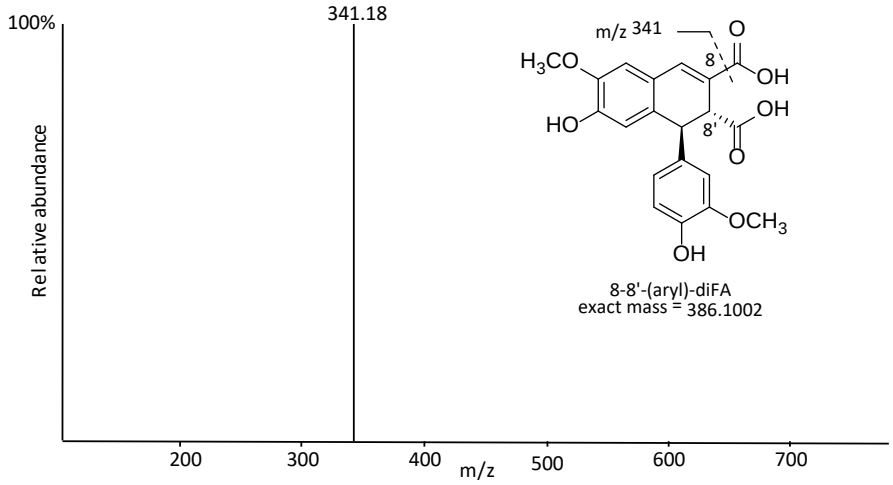


Figure A3: MS² spectrum for 8-8'-(aryl)-diFA

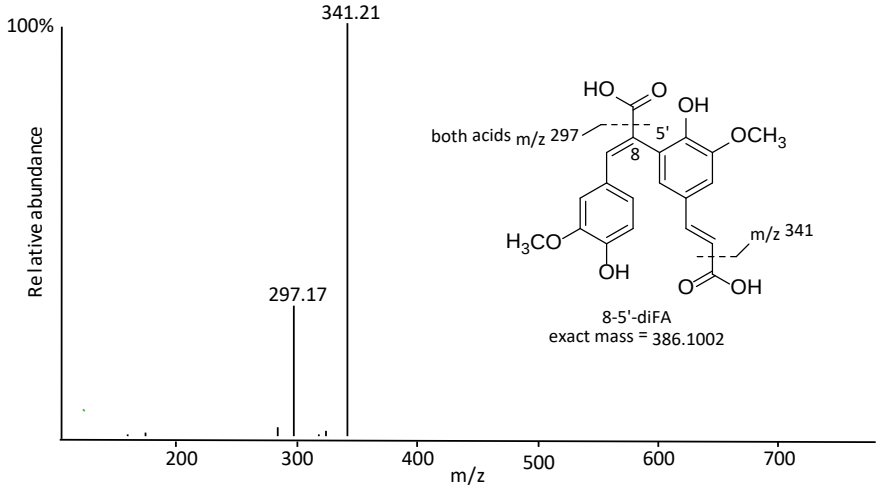


Figure A4: MS² spectrum for 8-5'-diFA

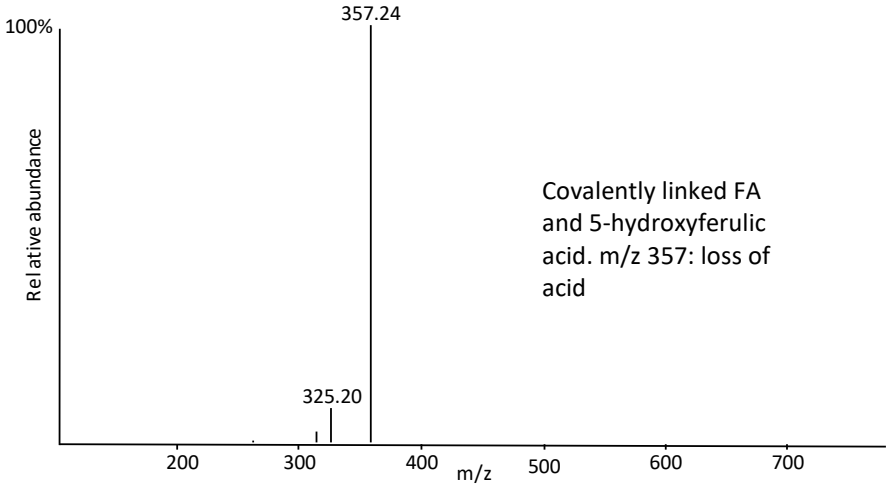


Figure A5: MS^2 spectrum of m/z = $401_{9.75\ min}$

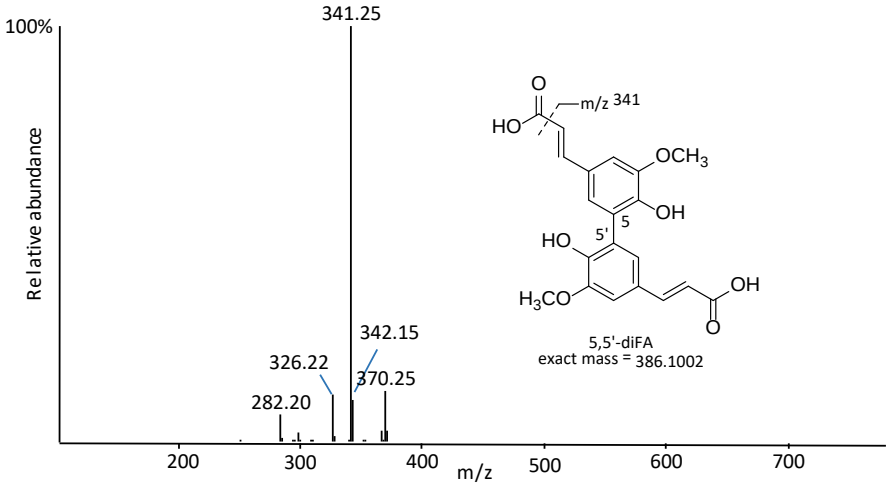


Figure A6: MS^2 spectrum of 5-5'-diFA

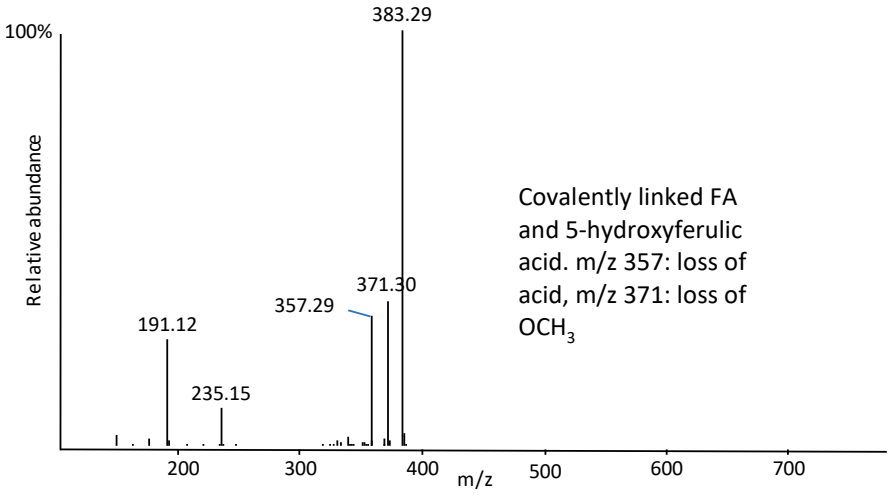


Figure A7: MS^2 spectrum of m/z = $401_{10.66\ min}$

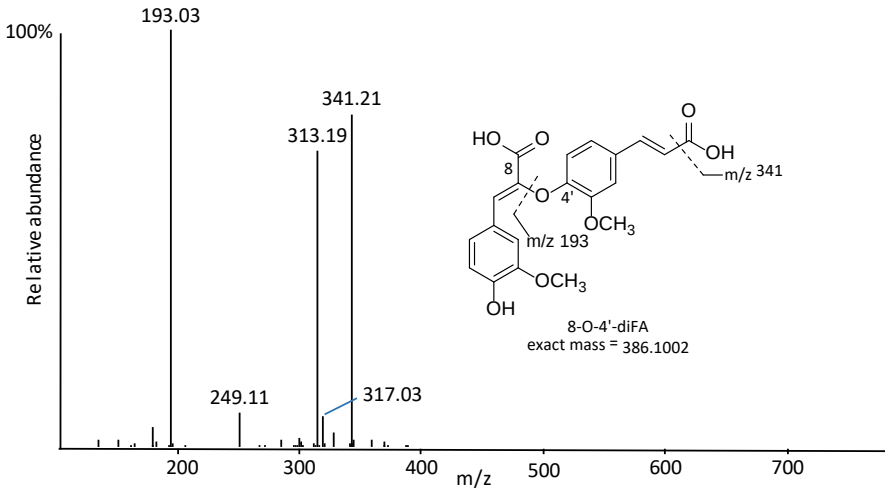


Figure A8: MS^2 spectrum of 8-O-4'-diFA

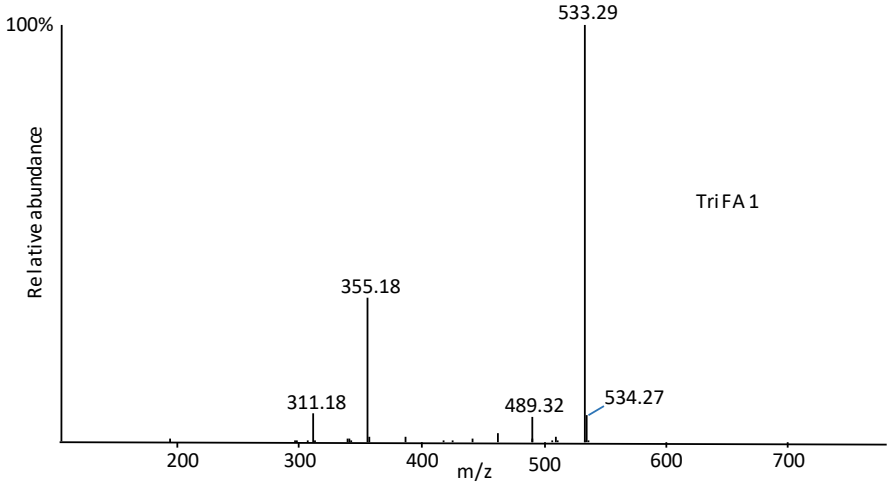


Figure A9: MS^2 spectrum of $triFA_{11.35\ min}$

Table A2: The release and standard deviation of FA and pCA as percentages of the bound content of all natural substrates

Substrate	Enzymes	Broth	AnFaeB1	AsFaeF1	AnFaeC5	AnidFAEC5	AsFaeC5	C1FaeA15	C1FaeA25	FoFae26	AsFaeE6	C1FaeB26	AnFaeA7	AnFaeJ9	AsFaeI13	ShFae113
SBP	FA	0.04	0.47	1.12	0.16	0.54	0.35	4.11	3.06	3.73	1.08	4.16	0.35	0.05	1.29	0.06
	Std. dev.$_{FA}$	0.00	0.01	0.04	0.00	0.08	0.00	0.10	0.05	0.04	0.36	0.10	0.01	0.00	0.10	0.00
	pCA	0.00	3.72	4.73	4.30	2.78	4.16	5.45	2.01	6.17	3.31	3.57	4.15	2.54	5.90	4.62
	Std. dev.$_{pCA}$	0.00	0.19	0.21	0.05	0.09	0.09	0.07	0.12	0.13	0.09	0.05	0.16	0.08	0.26	0.03
CFoligo	FA	8.43	28.84	17.73	21.34	32.91	18.64	73.37	41.05	39.00	13.89	31.18	84.21	16.23	23.21	19.16
	Std. dev.$_{FA}$	0.02	0.86	0.21	0.49	0.18	0.11	1.54	1.04	3.78	0.16	1.25	3.61	0.01	1.05	0.40
	pCA	0.00	30.36	5.55	9.06	13.70	8.03	20.96	39.11	40.01	8.79	19.84	17.67	5.92	8.76	11.34
	Std. dev.$_{pCA}$	0.00	0.25	0.08	0.10	0.10	0.08	0.01	0.21	2.03	0.10	0.67	0.73	0.06	0.47	0.25
WAX-i	FA	0.00	0.22	0.26	0.69	1.74	1.40	9.92	0.82	0.18	2.00	2.42	2.27	0.00	0.26	0.00
	Std. dev.$_{FA}$	0.00	0.01	0.004	0.04	0.41	0.03	0.28	0.02	0.01	0.03	0.14	0.29	0.00	0.01	0.00
	pCA	0.03	0.10	0.00	0.04	0.26	0.00	1.12	0.00	0.02	0.19	0.44	0.19	0.00	0.00	0.00
	Std. dev.$_{pCA}$	0.00	0.00	0.00	0.00	0.002	0.00	0.01	0.00	0.00	0.00002	0.01	0.01	0.00	0.00	0.00
Corn stover	FA	0.05	5.92	5.18	14.34	33.89	21.20	50.85	30.14	5.39	25.86	33.64	42.93	0.08	5.99	0.41
	Std. dev.$_{FA}$	0.00	0.02	0.24	0.70	0.03	0.28	1.35	0.75	0.07	0.25	0.65	0.39	0.00	0.09	0.01
	pCA	0.14	8.21	3.90	2.51	6.63	4.76	7.45	11.43	12.31	10.22	12.22	2.76	0.46	4.23	0.41
	Std. dev.$_{pCA}$	0.00	0.01	0.25	0.11	0.02	0.04	0.21	0.22	0.10	0.07	0.15	0.06	0.00	0.09	0.01
Wheat straw	FA	0.05	5.28	4.84	24.59	32.89	26.02	46.24	29.45	10.96	30.79	30.85	32.61	0.10	6.37	0.10
	Std. dev.$_{FA}$	0.002	0.29	0.17	1.38	0.92	0.46	0.34	0.54	0.25	0.71	0.02	1.18	0.0002	0.02	0.0003
	pCA	0.09	6.85	3.05	5.05	6.92	4.96	7.78	10.63	12.05	9.60	9.92	1.87	0.50	4.73	0.24
	Std. dev.$_{pCA}$	0.00	0.32	0.04	0.28	0.17	0.10	0.09	0.27	0.06	0.16	0.20	0.09	0.01	0.09	0.00
CSlignin	FA	0.13	2.67	8.00	7.20	16.41	0.13	29.03	29.09	9.07	29.44	25.48	28.66	0.82	12.22	1.16
	Std. dev.$_{FA}$	0.01	0.22	0.58	0.20	1.42	0.34	0.43	0.37	0.58	1.10	0.11	0.94	0.01	0.33	0.05
	pCA	0.17	3.75	1.28	1.21	3.51	2.01	3.45	7.62	5.95	6.26	5.66	0.83	0.25	2.16	0.03
	Std. dev.$_{pCA}$	0.00	0.24	0.01	0.02	0.13	0.12	0.01	0.22	0.09	0.19	0.05	0.03	0.01	0.09	0.00

Table A3: The release and standard deviation of diFAs and triFAs as percentages of the bound content of all natural substrates

Substrate	Enzymes / Compound	Broth	AnFaeB 1	AsFaeF 1	AnFaeC 5	AnidFAEC 5	AsFaeC 5	C1FaeA1 5	C1FaeA2 5	FoFae2 6	AsFeE 6	C1FaeB2 6	AnFaeA 7	AnFaeJ 9	AsFaeI 13	ShFae1 13
SBP	8-8’-furan-FA	0.00	0.00	0.00	0.00	**0.22**	0.00	**3.48**	**2.90**	**2.63**	0.00	**3.71**	0.00	0.00	0.00	0.00
	std. dev.$_{8\text{-}8'\text{-furan-FA}}$	0.00	0.00	0.00	0.00	**0.00**	0.00	**0.06**	**0.15**	**0.14**	0.00	**0.34**	0.00	0.00	0.00	0.00
	8-5'-FA	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	std. dev.$_{8\text{-}5'\text{-FA}}$	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	m/z = 401$_{9.75\ min}$	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	std. dev.$_{m/z\ 401}$	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	5-5'-FA	0.00	0.00	0.00	0.00	0.00	0.00	**1.48**	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	std. dev.$_{5\text{-}5'\text{-FA}}$	0.00	0.00	0.00	0.00	0.00	0.00	**0.09**	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	8-O-4'-FA	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	std. dev.$_{8\text{-}O\text{-}4'\text{-FA}}$	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	Tri-FA 1	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	std. dev.$_{TriFA\ 1}$	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	Tri-FA 2	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	std. dev.$_{TriFA\ 2}$	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
CFoligo	8-8’-furan-FA	**0.40**	**3.79**	0.00	**5.67**	**8.23**	**5.14**	**18.33**	**16.00**	**0.39**	**2.93**	**4.34**	**3.53**	**1.36**	**3.55**	0.00
	std. dev.$_{8\text{-}8'\text{-furan-FA}}$	**0.01**	**0.39**	0.00	**0.26**	**0.13**	**0.34**	**0.60**	**1.19**	**0.06**	**0.31**	**0.23**	**0.30**	**0.10**	**0.32**	0.00
	8-5'-FA	**0.28**	**0.73**	0.00	**2.48**	**2.32**	**0.59**	**20.11**	**8.79**	0.00	0.00	0.00	**0.63**	0.00	0.00	**3.27**
	std. dev.$_{8\text{-}5'\text{-FA}}$	**0.00**	**0.01**	0.00	**0.13**	**0.05**	**0.00**	**0.38**	**0.38**	0.00	0.00	0.00	**0.00**	0.00	0.00	**0.17**
	m/z = 401$_{9.75\ min}$	0.00	**4.48**	0.00	**3.33**	**15.83**	**6.41**	**87.07**	**23.14**	**11.46**	0.00	0.00	**38.07**	**0.50**	0.00	**1.69**
	std. dev.$_{m/z\ 401}$	0.00	**0.06**	0.00	**0.02**	**0.11**	**0.01**	**2.72**	**1.10**	**1.08**	0.00	0.00	**1.90**	**0.01**	0.00	**0.06**
	5-5'-FA	0.00	**0.00**	0.00	**7.57**	**22.84**	**9.11**	**90.77**	**43.54**	0.00	0.00	0.00	**82.35**	0.00	0.00	0.00
	std. dev.$_{5\text{-}5'\text{-FA}}$	0.00	**0.00**	0.00	**0.33**	**0.01**	**0.03**	**1.65**	**2.29**	0.00	0.00	0.00	**3.74**	0.00	0.00	0.00
	8-O-4'-FA	0.00	**0.64**	**2.03**	**3.91**	**8.95**	**5.29**	**12.51**	**11.15**	**0.65**	**0.23**	**0.34**	**0.10**	**1.43**	**0.90**	**1.75**
	std. dev.$_{8\text{-}O\text{-}4'\text{-FA}}$	0.00	**0.01**	**0.05**	**0.22**	**0.48**	**0.45**	**0.03**	**0.01**	**0.11**	**0.01**	**0.00**	**0.01**	**0.08**	**0.07**	**0.00**

	Tri-FA 1	0.00	0.00	**1.35**	0.00	0.00	0.00	**17.34**	**6.74**	0.00	0.00	0.00	**2.90**	0.00	**0.71**	0.00
	std. dev.$_{\text{TriFA 1}}$	0.00	0.00	**0.09**	0.00	0.00	0.00	**0.43**	**1.36**	0.00	0.00	0.00	**0.32**	0.00	**0.20**	0.00
	Tri-FA 2	0.00	0.00	0.00	0.00	0.00	0.00	**4.14**	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	std. dev.$_{\text{TriFA 2}}$	0.00	0.00	0.00	0.00	0.00	0.00	**0.01**	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
WAX-i	8-8’-furan-FA	0.00	**1.33**	**1.41**	**3.62**	**8.23**	**5.21**	**51.53**	**3.56**	**0.40**	**8.89**	**13.24**	**12.53**	**0.11**	**1.31**	**0.34**
	std. dev.$_{\text{8-8’-furan-FA}}$	0.00	**0.06**	**0.12**	**0.20**	**2.13**	**0.18**	**13.79**	**0.23**	**0.02**	**0.19**	**0.50**	**0.50**	**0.00**	**0.07**	**0.03**
	8-5'-FA	0.00	0.00	0.00	0.00	0.00	0.00	**4.15**	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	std. dev.$_{\text{8-5'-FA}}$	0.00	0.00	0.00	0.00	0.00	0.00	**0.23**	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	m/z = 401$_{\text{9.75 min}}$	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	std. dev.$_{\text{m/z 401}}$	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	5-5'-FA	0.00	0.00	0.00	0.00	**0.15**	0.00	**2.04**	0.00	0.00	0.00	0.00	**0.14**	0.00	0.00	0.00
	std. dev.$_{\text{5-5'-FA}}$	0.00	0.00	0.00	0.00	**0.03**	0.00	**0.09**	0.00	0.00	0.00	0.00	**0.03**	0.00	0.00	0.00
	8-O-4'-FA	0.00	0.00	0.00	0.00	0.00	0.00	**0.23**	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	std. dev.$_{\text{8-O-4'-FA}}$	0.00	0.00	0.00	0.00	0.00	0.00	**0.01**	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	Tri-FA 1	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	std. dev.$_{\text{TriFA 1}}$	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	Tri-FA 2	0.00	0.00	0.00	0.00	0.00	0.00	**0.10**	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	std. dev.$_{\text{TriFA 2}}$	0.00	0.00	0.00	0.00	0.00	0.00	**0.01**	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Corn stover	8-8’-furan-FA	**0.73**	**0.25**	0.00	**21.50**	**39.83**	**28.19**	**60.58**	**40.42**	**0.27**	0.00	**9.72**	**21.88**	0.00	0.00	**0.29**
	std. dev.$_{\text{8-8’-furan-FA}}$	**0.04**	**0.01**	0.00	**0.70**	**0.31**	**0.11**	**2.05**	**1.11**	**0.02**	0.00	**0.34**	**0.76**	0.00	0.00	**0.00**
	8-5'-FA	0.00	0.00	0.00	0.00	**27.81**	0.00	**78.47**	**48.48**	0.00	0.00	0.00	**76.06**	0.00	0.00	0.00
	std. dev.$_{\text{8-5'-FA}}$	0.00	0.00	0.00	0.00	**1.96**	0.00	**0.27**	**0.27**	0.00	0.00	0.00	**0.60**	0.00	0.00	0.00
	m/z = 401$_{\text{9.75 min}}$	0.00	**0.89**	**1.66**	**5.77**	**19.75**	**9.00**	**47.16**	**14.61**	0.00	**0.14**	**2.89**	**21.36**	**0.94**	**0.15**	**1.67**
	std. dev.$_{\text{m/z 401}}$	0.00	**0.01**	**0.07**	**0.57**	**0.11**	**0.33**	**1.64**	**0.55**	0.00	**0.01**	**0.07**	**0.50**	**0.03**	**0.01**	**0.04**
	5-5'-FA	0.00	0.00	0.00	**16.08**	**41.51**	**23.02**	**74.12**	**33.67**	0.00	0.00	0.00	**56.43**	0.00	0.00	0.00
	std. dev.$_{\text{5-5'-FA}}$	0.00	0.00	0.00	**0.20**	**0.75**	**1.45**	**3.26**	**0.11**	0.00	0.00	0.00	**1.93**	0.00	0.00	0.00

	8-O-4'-FA	0.00	0.00	0.00	**6.65**	**7.43**	**5.73**	**10.89**	**8.22**	0.00	0.00	**3.80**	**2.16**	0.00	0.00	0.00
	std. dev.$_{8\text{-}O\text{-}4'\text{-}FA}$	0.00	0.00	0.00	**0.53**	**0.03**	**0.20**	**0.50**	**0.15**	0.00	0.00	**0.28**	**0.07**	0.00	0.00	0.00
Wheat straw	8-8’-furan-FA	0.00	0.00	0.00	**3.64**	**36.45**	**28.80**	**50.71**	**36.19**	0.00	0.00	**10.59**	**9.71**	0.00	0.05	0.00
	std. dev.$_{8\text{-}8'\text{-}furan\text{-}FA}$	0.00	0.00	0.00	**1.11**	**0.77**	**0.06**	**0.92**	**0.41**	0.00	0.00	**0.12**	**1.62**	0.00	0.00	0.00
	8-5'-FA	0.00	0.00	0.00	0.00	0.00	0.00	**8.41**	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	std. dev.$_{8\text{-}5'\text{-}FA}$	0.00	0.00	0.00	0.00	0.00	0.00	**0.81**	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	m/z = 401$_{9.75\ min}$	0.00	0.00	0.00	**19.95**	**48.23**	**17.10**	**92.35**	**31.31**	0.00	0.00	0.00	**20.65**	0.00	0.00	0.00
	std. dev.$_{m/z\ 401}$	0.00	0.00	0.00	**0.34**	**3.04**	**0.71**	**0.10**	**1.20**	0.00	0.00	0.00	**1.32**	0.00	0.00	0.00
	5-5'-FA	0.00	0.00	0.00	**30.51**	**40.88**	**27.78**	**64.02**	**37.89**	0.00	0.00	0.00	**37.70**	0.00	0.00	0.00
	std. dev.$_{5\text{-}5'\text{-}FA}$	0.00	0.00	0.00	**0.00**	**2.52**	**0.38**	**0.54**	**0.58**	0.00	0.00	0.00	**1.81**	0.00	0.00	0.00
	8-O-4'-FA	0.00	0.00	0.00	0.00	**7.65**	0.00	**10.03**	**11.19**	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	std. dev.$_{8\text{-}O\text{-}4'\text{-}FA}$	0.00	0.00	0.00	0.00	**0.59**	0.00	**0.90**	**0.16**	0.00	0.00	0.00	0.00	0.00	0.00	0.00
CSlignin	8-8’-furan-FA	0.00	0.00	0.00	**0.86**	**1.75**	**1.09**	**2.75**	**3.05**	0.00	0.00	**0.42**	**1.21**	0.00	**0.14**	0.00
	std. dev.$_{8\text{-}8'\text{-}furan\text{-}FA}$	0.00	0.00	0.00	**0.06**	**0.12**	**0.06**	**0.13**	**0.12**	0.00	0.00	**0.01**	**0.03**	0.00	**0.01**	0.00
	8-5'-FA	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	std. dev.$_{8\text{-}5'\text{-}FA}$	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	m/z = 401$_{9.75\ min}$	0.00	0.00	0.00	0.00	**3.51**	**0.00**	**5.97**	**4.12**	0.00	0.00	0.00	**5.01**	0.00	0.00	0.00
	std. dev.$_{m/z\ 401}$	0.00	0.00	0.00	0.00	**0.17**	**0.00**	**0.36**	**0.02**	0.00	0.00	0.00	**0.21**	0.00	0.00	0.00
	5-5'-FA	0.00	0.00	0.00	**36.59**	**61.35**	**45.24**	**74.18**	**65.32**	0.00	0.00	0.00	**97.03**	0.00	0.00	0.00
	std. dev.$_{5\text{-}5'\text{-}FA}$	0.00	0.00	0.00	**3.41**	**1.58**	**0.95**	**7.52**	**0.19**	0.00	0.00	0.00	**3.45**	0.00	0.00	0.00

Table A4: Exact values for the release of m/z = $401_{10.66\ min}$

Enzymes	Broth	AnFaeB 1	AsFaeF 1	AnFaeC 5	AnidFAEC 5	AsFaeC 5	C1FaeA1 5	C1FaeA2 5	FoFae2 6	AsFeE 6	C1FaeB2 6	AnFaeA 7	AnFaeJ 9	AsFaeI 13	ShFae1 13
CFoligo	0.00	**0.37**	0.00	**1.48**	**5.92**	**3.57**	**13.33**	**11.14**	**1.98**	0.00	**1.36**	**6.47**	**0.25**	0.00	0.00
Std. dev.$_{CFoligo}$	0.00	**0.00**	0.00	**0.03**	**0.03**	**0.23**	**0.03**	**0.78**	**0.00**	0.00	**0.06**	**0.90**	**0.01**	0.00	0.00
Corn stover	0.00	**0.22**	0.00	**1.67**	**3.15**	**2.59**	**4.18**	**3.53**	**0.19**	**0.16**	**1.97**	**0.96**	0.00	**0.17**	0.00
Std. Dev.$_{corn\ stover}$	0.00	**0.02**	0.00	**0.11**	**0.07**	**0.06**	**0.22**	**0.19**	**0.01**	**0.01**	**0.16**	**0.03**	0.00	**0.00**	0.00
Wheat straw	0.00	**0.05**	0.00	**2.24**	**2.48**	**2.12**	**3.23**	**2.35**	**0.04**	0.00	**1.56**	**1.97**	**0.05**	**0.10**	0.00
Std. Dev.$_{wheat\ straw}$	0.00	**0.01**	0.00	**0.04**	**0.08**	**0.03**	**0.13**	**0.06**	**0.00**	0.00	**0.02**	**0.09**	**0.00**	**0.00**	0.00
CSlignin	0.00	**0.17**	0.00	**1.19**	**2.20**	**1.94**	**1.98**	**2.49**	**0.25**	**0.18**	**1.41**	**0.35**	0.00	0.00	0.00
Std. Dev.$_{CSlignin}$	0.00	**0.01**	0.00	**0.06**	**0.17**	**0.07**	**0.16**	**0.19**	**0.00**	**0.00**	**0.01**	**0.02**	0.00	0.00	0.00